LECTURE NOTES ON SEMINAR ON SCHEDULE DESIGN

for Psychological and Educational Survey Researches

Edited by
D. Dutta Roy

Proceedings of Seminar
December 8, 2006

PSYCHOLOGY RESEARCH UNIT
INDIAN STATISTICAL INSTITUTE
203,B.T.ROAD,KOLKATA-700108

LECTURE NOTES ON SEMINAR ON SCHEDULE DESIGN

for Psychological and Educational Survey Researches

Edited by
D. Dutta Roy

Proceedings of Seminar
December 8, 2006

PSYCHOLOGY RESEARCH UNIT
INDIAN STATISTICAL INSTITUTE
203,B.T.ROAD,KOLKATA-700108

for Acc
RS.

Lecture Notes on Seminar on Schedule design for Psychological and Educational Survey Researches

Proceedings of Seminar
December 8, 2006

Edited by
D. Dutta Roy

PSYCHOLOGY RESEARCH UNIT
INDIAN STATISTICAL INSTITUTE
203, B.T. ROAD, KOLKATA – 700 108

CONTENTS

Preface

Though schedule design is part and parcel of large-scale sample survey in Psychological and Educational researches, several principles of schedule design are not properly focused in the research methodology books of Psychology and Education. On this backdrop, Psychology Research Unit of the Indian Statistical Institute organized one seminar titled " Seminar on Schedule design for Psychological and Educational Survey Researches" dated 8.12.06. The target participants were the Research fellows or scholars and Project workers and faculty members of the Universities and the Institutes. All the participants received participation certificates at the valedictory session.

Current volume contains lecture notes presented in the seminar. It has three parts (a) Exploring variables (b) Scaling and Schedule design and (c) Pilot testing. In part 1, Dr. Prashanta Pathak focused attention on how to identify the explanatory variables specially in analytical researches. He also made difference between explanatory and causal variables. Dr. Jayanti Basu highlighted several ethical issues involved in interview for schedule design. Dr. Debjani Sengupta using pictorial representations explained different structures of enquiry.

In Part II, Dr. Susmita Mukhopadhyay meticulously explained how to scale the different types of variables and how to determine their relation using different statistical models. She also demonstrated different kinds of schedules used in Psychological and Educational researches. Dr. Pulakesh Maity explained in details how to overcome from different types of errors in survey and in schedule design. Professor Arijit Choudhury focused attention on weighting adjustment and imputation technique in handling non response errors. Dr. D. Dutta Roy discussed different steps in

schedule design and principles to manage dissimilarity in scaling properties of schedule after collection of data.

In Part III, Professor S.P. Mukherjee emphasized on different reasons for pilot testing so that final schedule can provide more accurate, reliable and valid information about respondents. Professor Prafulla Chakraborty described in details importance of pilot testing in social science researches.

I am grateful to all the authors who made the publication of volume possible. I sincerely acknowledge the administrative and financial supports of Professor Shankar Pal, the Director of the Indian Statistical Institute, Professor Tarun Kabiraj, Professor-in-charge, Social Sciences Division and Dr. Anjali Ghosh, Head of the Psychology Research unit of the Indian Statistical Institute to organize the workshop and to prepare this volume.

I am grateful to Professor Manjula Mukerjee, Director, Indian Institute of Psychometry, Kolkata, Ex-Head, Psychology Research Unit, ISI, Kolkata for her speech on schedule design and to Mr. B.K. Giri , Deputy Director General, DPD, National Sample Survey Organization for his key note address on schedule design for National sample survey on health and education in the inaugural session. At the valedictory session, Dr. Rumki Gupta, my colleague, gave Vote of thanks. I am thankful to her.

Our research scholars Ms. Rita Karmakar, Ms. Fouzia Alsabah Shaikh and project assistant Ms. Rituparna Basak tried their best to make the seminar success.

The seminar can not be organized without constant support of the staffs of our unit. Therefore I am very much grateful to Shri Shyamal Chatterjee, (Section officer) , Mr.Nandagopal Chakraborty, Mr. Ardhendu Bhattacharya, Dr. (Mrs.) Himani Bhattacharya, Mr. Basanta Santra, Mr. Ramlal Prasad and Mrs. Shanti Hela.

Hope, Lecture notes collected in this volume will stimulate researchers in Psychology and Education to develop interest in designing different schedules for their research pursuits.

D. Dutta Roy
Editor and Convener of the Workshop

ANNOUNCEMENT

Seminar on

Schedule design for Psychological and Educational Survey Researches

Date : December 8, 2006
Time: 10 am to 6 pm
Organized by : Psychology Research Unit , Indian Statistical Institute, Kolkata.
Following issues will be examined and discussed **in the seminar.** 1. Identification of explanatory variables
2. Interview strategies and Ethical issues 3. Accounting of intervening variables 4. Scaling and Schedule design 5. Non response and the reasons thereof 6. Handling dissimilarity in scaling properties across different response categories of schedule 7. Pilot testing of schedule
Who can apply: Faculty members, research fellows, project workers of different institutes and universities who are involved in Psychological and Educational Survey researches. Seats are limited.
Seminar Fees: There is no charge for training or fees for registration in the Seminar
Financial Support : No financial support will be provided to the participants for traveling/ boarding and lodging. Participants will have to make their own arrangement for boarding and lodging .
To apply : Registration form is available in the seminar website. Please fill up the form and send it through proper sponsorship channel to Convenor, "Seminar on Schedule design for Psychological and Educational Survey Researches", Psychology Research Unit, Indian Statistical Institute,203, Barrackpore Trunk Road, Kolkata 700 108,FAX: 033-2577-6925,/ 6033,Telephone(s):25753450/25753454
Last Date : Last date of receiving application is 24th November, 2006
E-mail contact: psy@isical.ac.in
Website : http://www.isical.ac.in/~psy/sem/Sem.htm

D. Dutta Roy
(Convenor)

REGISTRATION FORM
Seminar on Schedule design for Psychological and Educational Survey Researches
December 8, 2006

Psychology Research Unit , Indian Statistical Institute
203 Brrackpore Trunk Road, Kolkata - 700 108
E-mail: psy@isical.ac.in

TO BE FORWARDED THROUGH PROPER SPONSORING CHANNEL

1. Name (in Block Letters): Mr./Ms.:
2. Designation :
3. Office Address :
4. Home Address :
5. Telephone no. :
6. Fax/e-mail Address (if any) :
7. Age (in completed years) :
8. Qualification :
9. Brief description of your uses of Schedules in current or prior researches (Use separate page):
10. Expectation (and/or requirements) from the Seminar :
11. Signature of candidate with date :

ENDORSEMENT

This office is sponsoring the candidature of Mr./Ms. _____(Designation) _____

Working at _____ for the above Seminar.

Office Seal _____ Signature of the

Sponsoring Authority (Name Designation:)

Date : Application dead-line : 24th Nov., 2006 (This form may be copied and circulated widely. You can use separate sheet as annexure)

Application form is available in the website:
Website : http://www.isical.ac.in/~psy/sem/Sem.htm
http://www.isical.ac.in/~psy/sem/schsem.pdf

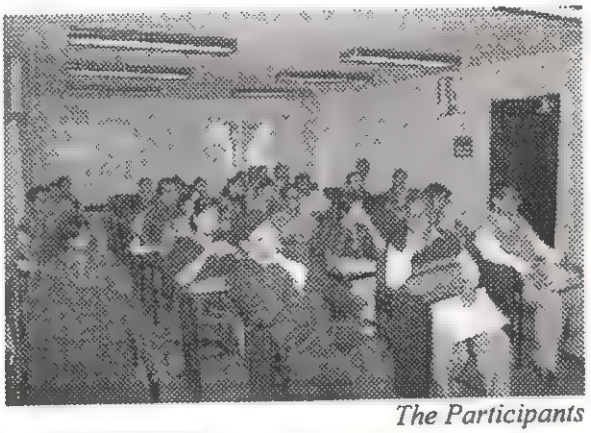
Program Schedule

Time	Topics/Events	Resource Persons
9:30	Registration : Verification and Kit distribution	
10:00-10:45 AM	INAUGURATION	Professor Tarun Kabiraj,(*Professor-in-charge , Social Sciences Division.)*
	Welcome Address	Dr.Anjali Ghosh *(Head, Psychology Research Unit)*: Professor Manjula Mukerjee, *Director, Indian Institute of Psychometry, Kolkata, Ex-Head, Psychology Research Unit, ISI, Kolkata.*
	Key Note Address on "Schedule design for National Sample Survey on Health & Education"	Shri B.K. Giri, *Deputy Director General, DPD, National Sample Survey Organization.*
10:45-11:15	IDENTIFICATION OF EXPLANATORY VARIABLES	Dr. Prasanta Pathak, *Population Studies Unit, Indian Statistical Institute, Kolkata.*
11:15-11:30	TEA BREAK	
11:30-12:00	IDENTIFICATION OF EXPLANATORY VARIABLES	*Professor Debjani Sengupta, Department of Education, Calcutta University*
12:15-12:45	INTERVIEW STRATEGY AND ETHICS: CONTEMPLATING SOME REAL LIFE ISSUES	*Dr. Jayanti Basu, Head, Department of Applied Psychology, Calcutta University*
13:00-13:30	ROLE OF PILOT SURVEY IN SOCIAL SCIENCE: RESEARCH METHOD	*Professor Prafulla Chakrabarti, Director of Research, SERI, Mohana, 5, New Raipur, Kolkata*
13:45-14:30	LUNCH BREAK	
14:30-15:00	PILOT TESTING OF SCHEDULE	*Professor S. P. Mukherjee, Centenery Professor, Department of Applied Statistics, Calcutta University, Kolkata*

15:15-15:45	NON-RESPONSE AND REASONS THEREOF	*Professor Pulakesh Maity, Economic Research Unit ,*
15:45-16:15	NON-RESPONSE AND REASONS THEREOF	*Professor Arijit Choudhury, Visiting Professor, Applied Stat. Unit, Indian Statistical Institute, Kolkata*
16:30 - 16:45	COFFEE BREAK	
16:45-17:15	SCALING IN SCHEDULE DESIGN AND ACCOUNTING OF INTERVENING VARIABLES	*Dr. Susmita Mukhopadhyay, Visiting Faculty, VGSOM, IIT., Kharagpur.*
17:30-18:00	MANAGING DISSIMILARITY IN SCALING PROPERTIES OF SCHEDULE	*Dr. D. Dutta Roy, Psychology Research Unit, Indian Statistical Institute, Kolkata*
18:00-18:30	VALEDICTORY SESSION AND DISTRIBUTION OF CERTIFICATES	*Vote of Thanks by: Dr. Rumki Gupta (Psychology Research Unit, ISI, Kolkata)*

Some Memorable Events

Dr. Anjali Ghosh, Mr. B.K. Giri, Professor Tarun Kabiraj and Professor Manjula Mukerjee *at the inaugural Session*



The Participants

Distribution of Participation certificate at the Valedictory Session

IDENTIFICATION OF EXPLANATORY VARIABLES

Prasanta Pathak
Population Studies Unit
Indian Statistical Institute

Among three major types of research studies namely, descriptive studies, analytical studies and exploratory studies, the ones where identification of explanatory variables becomes essential are of second type. The first type of studies is generally done for situational analysis, which may ultimately become useful in identification of explanatory variables. The third type of studies is generally done when existing knowledge about the studied variables is quite insufficient and hence identification of explanatory variables is rather difficult.

In any study, the objectives play the main role in identification of variables. In fact, the objectives are decided based on an in-depth problem analysis. Existing literature and available information are made use of in this problem analysis. The factors that are associated with the problem are identified, followed by identification of measurable variables that describe the factors best. Defining the variables in unambiguous terms and with highest objectivity is very essential as in absence of that information collection may face various measurement and interpretational problems. Among these variables, the dependent ones are chosen based on earlier studies and established theories. These earlier studies might be descriptive or analytical or exploratory. Earlier studies also help in

identifying the explanatory variables. Additionally, some new explanatory variables may be taken into consideration for enhancing the scope of a research study. All these variables basically correspond to different questions in a schedule if its values are collected through survey.

Often, explanatory variables are misinterpreted as causal variables. For a given dependent variable, explanatory variables are those that are highly associated with it. Establishing that an explanatory variable is also a causal variable is not an easy task as it involves use of an appropriate statistical design of experiment and/or use of a sophisticated statistical method of data analysis. There are various statistical methods of establishing association between any two variables. Methods of correlation analyses and contingent table analyses are the most common ones. High degrees of correlation among the explanatory variables should alert the researcher while choosing the explanatory variables. The explanatory variables that are highly associated might have a number of redundant variables. The researcher may keep only one or a few of them, which are most highly associated with the dependent variable. High association among some chosen explanatory variables sometimes misguides a researcher, making him/her think that one or more of them are redundant. Existence of one or more confounding variables, having high association with the chosen explanatory variables, sometimes results in such high association. Thorough problem analysis helps in identifying the confounding variables and thus helps in avoiding wrong classification of some explanatory variables as redundant variables. The points discussed above are explained below for a particular objective of an analytical study.

An Example:

If the objective of the study is to identify factors that influence drop out rate, the steps for identifying explanatory variables are broadly the following. It should, however, be noted that the dependent variable here is the rate of drop out. The variation in it is to be explained by a set of appropriately chosen explanatory variables.

The steps:

1. Define well the dependent variable to avoid any kind of ambiguity,

   e.g. (1) Is it drop out in class I or drop out in the course of moving from primary to middle level education? Or, (2) Is it for all or only for SC/ST students?

2. Do sufficient literature search to identify possible influencing factors.

3. Identify factors that are relevant in the context of your study. These may include factors that were not identified earlier.

4. Find out measuring variables for the identified factors. These could be nominal, ordinal, ratio or interval variables. Attempt should be made to

define variables as objectively as possible. The more is the level of quantification the more is the objectivity,

e.g. information on regularity of English classes may be collected as poor, reasonable, good, very good or may be collected as total scheduled classes in the last week and number of classes taken in the last week. The latter in appropriately combined form (say, per cent scheduled classes that took place) measures regularity more objectively.

5. Ensure that

(i) Every variable has been defined in the best possible way,

e.g. "Absenteeism in English class" is poorly defined as its interpretation could be more than one. Better variable is "Percentage absent in English class".

(ii) Redundant variables have not been included,

e.g. if the above mentioned better variable is included in the list of explanatory variables inclusion of "Percentage present in English class" or "Interest of students in attending English class" becomes redundant (unless the information is collected for cross verification).

(iii) Confounding variables have been appropriately taken care of,

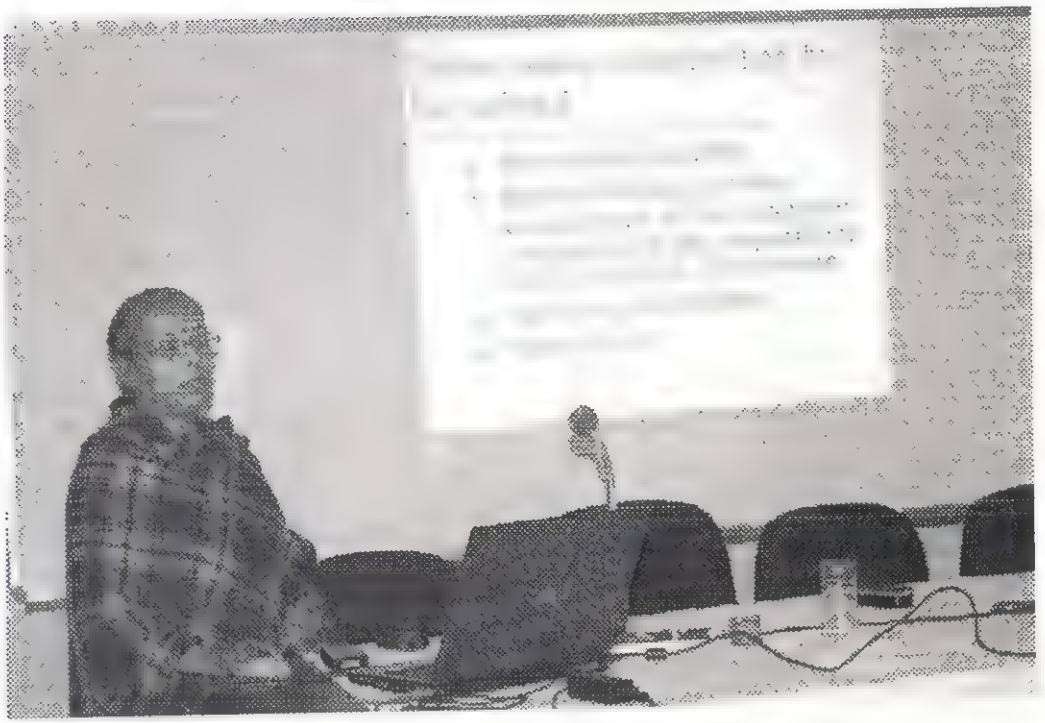
e.g. with increasing monthly income of guardians drop out rate is likely to fall. Again, with increasing monthly income guardians are likely to get their children trained by private tutors. However, it might not be an acceptable hypothesis that the drop out rate could be decreased by engaging more and more private tutors.

Importance of a thorough problem analysis has been already emphasized. It immensely helps in identifying those explanatory variables that are at the root of a given research problem, e.g. *irregularity in taking classes by concerned teachers might be a result of the difficulties of the teachers to reach the school in time due to non-availability of convenient transportation facilities. A village school, poorly connected with the nearby urban or semi-urban areas might experience greater drop out rate.* Systems theory based understanding of the research problem is also useful for selecting without fail the input, the process and the environmental explanatory variables. The input ones are those which are basic requirements for organizing a class, *e.g. a teacher, a class room, a black board* and so on. The process ones are those, without which the students do not get educated, *e.g. timely attendance by students, timely arrival by a teacher in the class, delivery of*

lecture by the teacher in the class as per schedule and so on. The environmental ones are those which act favourably or unfavourably in provision of input and/or in running the classes, *e.g. availability of teachers' training facility, availability of transportation facilities and so on.* Such systems theory based selection of variables helps research findings to be much more useful for effective implementation under various developmental programmes.

Interview strategy and ethics: Contemplating some real life issues

Jayanti Basu
Department of Applied Psychology
Calcutta university



By Interview strategy I mean the deliberate consideration of the predefined approach to interview process per se, as well as the inevitable 'accidents' and unforeseen awkward situations arising in most real life interviews. The solution to the latter may not be pre-designed, but entails an orientation and approach to dialogue with another person. With greater recognition of relativity and reflexivity in qualitative methodology, it is probably wiser to begin with admission of biases, prejudices, exploitation and blindness, rather than feigning neutrality. In this sense, any discussion on interview and exploratory survey needs to adequately highlight these subtler nuances of dyadic dynamics.

At the same time, one must not forget that every research and interview thereof entails an element of creativity, and hence a dose of serendipity also, as creative moments cannot be predicted. Of course we have a pre-designed schedule, but in most cases the brilliance of the results depends not on the predetermined scheme, but on the insight and modifications of the items during the course of the interview. To handle this aspect

properly, one needs to have a strong theoretical base of the subject matter as well as knowledge of the criteria of good interview.

The word 'interview' has been interpreted in various ways. Kvale (1996) regards interview as a way of bringing together the multiple views of people. Barbour & Schostak (2005) views interview as the space between the views, not the views themselves, but rather as the negative condition under which people may express their views to each other and to themselves. In the latter approach, interviewing is an experiment in the sense Rorscahch called inkblot testing an experiment.

Some of the key issues to be remembered in interview situations are: the power notion embedded in the process, the mutual social position, values, trust, meaning of words, interpretation and finally the uncertainty implied in any human interaction.

In an interview the following steps have to be designed, and the specific strategy for each needs be decided and loosely defined. I emphasize the word 'loosely', because, in most psychologically sensitive and open ended interview, the strategy actually changes from case to case. The nodes at where the strategy has to be defined are as follow:

1. Access to people
2. Range of concepts of discourse, or theoretical perspectives
3. The problem profile – its historical and social nuances, physical aspects and legal implications, if any
4. Recording the data
5. Representation of the experience of the research process and the experience of the subject of research
6. Analytic proceeding
7. Writing up

Some of the major difficulties in real life interview that cannot be truly taken care of, specifically in a predetermined way, are: the reactions to questions which, unknowingly have hurt the interviewee, socially or emotionally; unforeseen outcomes – good or bad or mixed; transference-countertransference issues and biases greeted from it; and finally the emotional reactions and change in self concept of the interviewer during and after the process (this involves development of the interviewer as well).

The ethical issues need to take into account all these features. The usual techniques of ethics are:

1. Consent form : Consent form includes
   - The purposes and intended use of the research
   - The expected duration of the interview or discussion and the nature of any incentive
   - The right of the respondent to decline to participate
   - Control over access to data collected (including how interview transcripts are archived or survey data is made available)
   - Contact details, if at any point of time, the respondent have any question, comment or complaint.

2. Confidentiality and anonymity : Success of an interview depends on the trust build up between the interviewer and interviewee. Particularly when a person is working with personal emotions and / or sensitive socio-political attitudes, the right to confidentiality and anonymity is a sine qua non of research ethics.

3. Prepublication access : Despite all efforts at anonymity of the informant, some data may be leaked to the disadvantage of the informant. Prepublication access refers to the informant's conscious agreement to the full form of presentation to be.

4. Ethical guidelines to be framed : Before a research is conducted, all possible aspects pertaining to the respondent's rights and consequences should be considered, and appropriate ethical guidelines need to be prepared. These include informed consent format, the language to be used, the personnel involved.

5. Ethical committees : Ethical committees need to be structured for supervising the proposals and treatment of research data. The constitution of ethical committees and specific roles differ from one country to another. Although there may be a controversy about the supervising role of such committees, most researchers feel that prepublication restriction and legal sanction is better to handle than post publication legal complication.

6. Situated ethics : Situated ethics is a view of applied ethics in which abstract standards from a culture or theory are considered to be far less important than the immediate ongoing processes in which one is personally and physically involved, e.g. climate, ecosystem, etc. This refers to the general sensitivity to ethical commitment so that the researcher may act as one's own guide in unforeseen and unique research situations.

However, in real context the meaning of these words often get confounded. There are works where anonymity is not recommended, for example in research on policy making. Here informed consent is more important. There are also a number of controversies regarding prepublication access. One is the gap between what one should do and the feasibility of practicing it in real life. Also, all informants may not be able to understand

the technicalities involved in the publication. Questions remain as to whether it is scientifically proper that participants be allowed to comment and add to the published matter? Furthermore, there are views asserting that it is not desirable for the researcher to be ethically neutral, when action research is being conducted.

It may be stated that there is no final and unanimous ethical standard for all researches. As Soltis (1990) pointed out, there are three different perspectives of the researcher, profession and public, and the inherent dilemma will remain forever. It is perhaps the extensive experience and self critical view only that guides the researcher in real life situation.

References

Barbour, R. S. & Schostak, J. (2005) Interviewing and focus groups. In B. Somekh & C. Lewin (Eds.) Research methods in the Social sciences . Vistaar. New Delhi. Pp. 41-48.

Kvale, S. (1996) Intwerviews: An introduction to qualitative research interviewing. Sage: New Delhi.

Soltis, J. F. (1990) The ethics of qualitative research. In E. W. Eisner & A. Peshkin (Eds.). Qualitative inquiry in education : The continuing debate. Columbia University : New York. Pp 247-257.

Structuring Enquiry

Debjani Sengupta
Department of Education
Calcutta University

From a Social Scientist's perspective, the purpose of research may be

- exploration of some phenomenon / issues / events
- description with all the details answering the what , when, how or other questions
- explanation of why things are happening in a particular fashion.

Taking the issue of women's access to higher education, one may explore the rate of enrollment of women students in higher education institutions. She may also describe the trend of enrollment over a span of time or within different subgroups. She may also identify the factors which can explain a particular trend of enrollment that she has observed in case of female students . It goes without saying that the above-mentioned purposes are not mutually exclusive of each other, and a good academic research must always aim for explanation with clarity.

The approach to explanation may be ideographic or homothetic. In ideographic explanation, the researcher is interested in explaining a single case thoroughly. For this one needs to take account of all possible factors / independent variables that may contribute to the issue. The approach is to study in depth or single case in minute details

so that none of the contributing variables remain unexplored. It helps to explain the research issue deterministically which, however, holds true for the particular case in consideration only. In contrast, nomothetic explanation identifies relatively fewer factors that are applicable for all other similar cases in general.

Casual explanation establishes that the phenomena Y (dependent variable) is affected by factor X (independent variable). It needs to be mentioned that any association between the DV and IV does not mean that they are casually associated. A casual relation is established only when it fulfils several criteria like

1) the variables are correlated,
2) the cause takes place before the effect, and
3) the variables are non spurious.

1. Correlation

Unless there is some associative relationship, one can not be certain that existence of one variable explains the occurrence of the other variable. Correlation, as it was stated earlier, is a necessary condition of causality but not a sufficient condition.

Variables may be associated in different ways depending on the degree of complexity of association.

Case I.

Direct one- to- one association

Preposition : women's representation is fewer in comparison to men in highly paid Jobs.

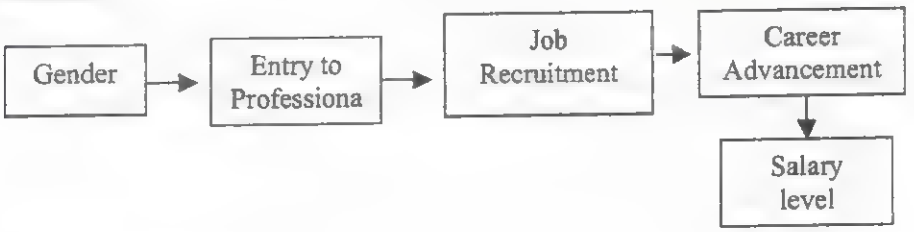
Relationship

Case II.

Indirect association bound in a causal chain

Preposition : Lesser no. of women are enrolled in professional causes that lead to highly paid jobs and ample professional advancement .

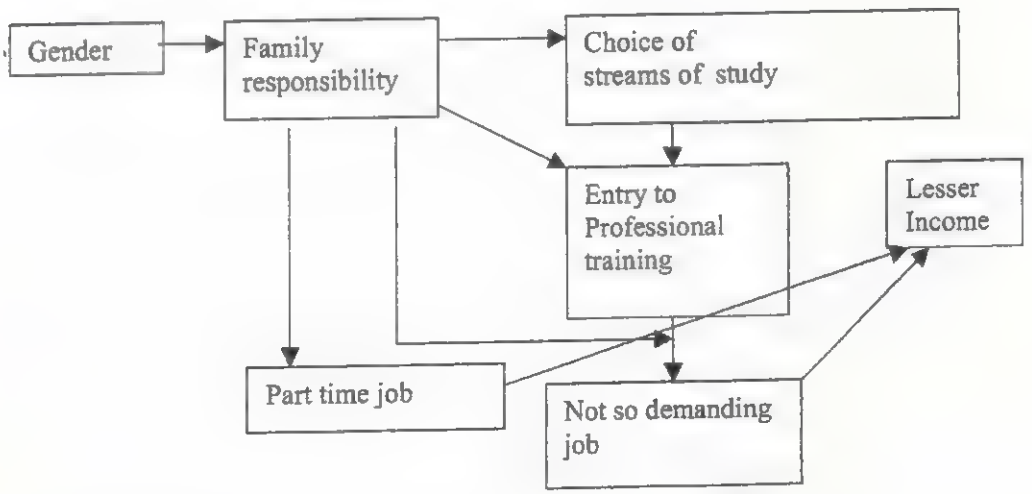
Relation



Case III

Complex association

Preposition : Social norm expect women to take care of the child and the aged which restricts their entry to selected professional courses. Both of these factors dictates women to take less demanding jobs or part time assignments leading to lesser income than males.



2. Time order
   A causal relationship can not exist unless the cause proceeds the effect. As in our previously cited example, a highly specialized professional degree is the pre-condition for getting a highly paid job.

3. Non spurious association

The third condition entails that the relationship must be genuine implying that the association dose not occur due to the presence of a third variable that has not been taken into account. As for example, at one time it was a popular belief that in general black people possess lower IQ than their white counter part. The actual fact remains that the content of the Intelligence tests that were used to assess Intelligence were heavily loaded with items that favored the life style and the culture of the white people. Sheer lack of familiarity with the content resulted in lower IQ measure of the black people.

Identifying the explanatory variables :

The following questions may help the researcher to clarify and substantiate the choice of variables:

- What am I trying to explain ?
- What are the possible causes ?
- Which causes will I explore ?
- What would be the possible mechanism to assess those variables ?

Exploring the following sources may help in identifying the variables :

- Previous Research : A thorough group of previously done research work and related theories on the same topic or allied topics often helps the researcher to obtain proper perspective for identifying the variables.

- Empirical findings : Researcher may sometimes observe the co-variance of two variables . In everyday life which suggests the possibility of a causal relationship. In the previously cited example, the researcher may often come across the fact that employed women share more family responsibility than the employed male folk which led her to work on an association between promotional opportunity and family responsibility.

- Talking to informants :

  Talking to people who are well formed about the issue or who, by any chance, are directly or indirectly associated with the issue may help to develop proper insight for identifying the variables. As for example, talking employers / professors of the training courses or to family members of some employed women as well as the women themselves may help to identify the appropriate explanatory variables.

This paper, in brief, offers some of the suggestions to get a perspective for conducting explanatory research. But it is , to be kept in mind all that the researches in education that belong to explanatory tradition e.g. some branches of historical research do not follow all these conditions particularly when the orientation is entirely ideographic.

However, the researcher must keep in mind that the explanations in social science are much more probabilistic than deterministic. We can only improve our probabilistic explanation by specifying the conditions under which the association holds true. Findings cannot invariable be true in all situations. The researcher must keep this limitation in mind while explaining relationship between the variables.

ACCOUNTING FOR VARIABLES, SCALING AND SCHEDULE DESIGNING FOR INTERVIEWS.

Dr. Susmita Mukhopadhyay
Visiting Faculty
Vinod Gupta School of Management
IIT, Kharagpur-721302

Accounting for variables, scaling and schedule designing for Interviews.

Dr. Susmita Mukhopadhyay
Visiting Faculty
Vinod Gupta School of Management
IIT, Kharagpur-721302

An INTERVIEW is a data-collection technique that involves oral questioning of respondents, either individually or as a group. An interview is a two-way purposeful conversation initiated by an interviewer to obtain information that is relevant to some research purpose. The participants are typically strangers and the topics and pattern of discussion are dictated by the interviewer. Interviewing is perhaps the most ubiquitous method of obtaining information from people. Interviews are ordinarily quite direct and a great deal of information is generally got from respondents by direct questioning (Emory, 1980; Kerlinger, 1980).

Answers to the questions posed during an interview can be recorded by writing them down (either during the interview itself or immediately after the interview) or by tape-recording the responses, or by a combination of both.

Interviews can be conducted with varying degrees of flexibility. The two extremes, high and low degree of flexibility, are described below:

• **High degree of flexibility**:

For example:

When studying sensitive issues such as teenage pregnancy and abortions, the investigator may use a list of topics rather than fixed questions. The investigator should have an additional list of topics ready when the respondent falls silent. The sequence of topics should be determined by the flow of discussion. It is often possible to come back to a topic discussed earlier in a later stage of the interview.

The unstructured or loosely structured method of asking questions can be used for interviewing individuals as well as groups of key informants. A flexible method of interviewing is useful if a researcher has as yet little understanding of the problem or situation he is investigating, or if the topic is sensitive. It is frequently applied in exploratory studies. The instrument used may be called an interview guide or **Interview schedule.**

- A schedule is filled in by the interviewer and is never mailed to the respondents
- It is generally used where the survey is to be conducted of a relatively small geographic area
- It can be used even when the respondents are illiterate
- In it the wording is not in the form of question
- In its designing the convenience of the investigator in handling it in the field should be the main consideration.

• **Low degree of flexibility:**

Less flexible methods of interviewing are useful when the researcher is relatively knowledgeable about expected answers or when the number of respondents being interviewed is relatively large.
For example:

After a number of observations on the (hygienic) behaviour of women drawing water at a well and some key informant interviews on the use and maintenance of the wells, one may conduct a larger survey on water use and satisfaction with the quantity and quality of the water.

Then **Questionnaires** may be used with a fixed list of questions in a standard sequence, which have mainly fixed or pre-categorized answers.

Questionnaire is

- Filled up by the respondent
- Is generally used when the field of enquiry is large
- Cannot be used where the respondents are illiterate
- In it wording is in the form of question
- In its designing knowledge, convenience and mood of the researcher should be the main consideration.

Flexible techniques, such as loosely structured interviews using open-ended questions is also called QUALITATIVE research techniques. They produce qualitative data that is often recorded in narrative form. QUALITATIVE RESEARCH TECHNIQUES involve the identification and exploration of a number of often mutually related variables that give INSIGHT in human behaviour (motivations, opinions, attitudes), in the nature and causes of certain problems and in the consequences of the problems for those affected. 'Why', 'What' and 'How' are important questions.

Structured questionnaires that enable the researcher to quantify pre- or post-categorized answers to questions are an example of QUANTITATIVE research techniques. The answers to questions can be counted and expressed numerically. QUANTITATIVE RESEARCH TECHNIQUES are used to QUANTIFY the size, distribution, and association of certain variables in a study population. 'How many?' 'How often?' and 'How significant?' are important questions.

Both qualitative and quantitative research techniques are often used within a single study.

For example:

It has been observed in country X that children between 1 and 2-1/2 years, who have already started to eat independently, have unsatisfactory food intake once they fall ill. A study could be designed to address this problem, containing the following stages:

Focus group discussions (FGDs) with 2 to 5 groups of mothers or in-depth interviews with 10 - 20 mothers, to find out whether they change the feeding practices for children in this age group when they suffer from (various) illnesses and how mothers deal with children who have no appetite when they are sick (exploratory study); A cross-sectional survey, testing the relevant findings of the exploratory study on a larger scale; and FGDs with women in the study area to discuss findings and possible questions arising from the survey and to develop possible solutions for problems detected.

In this example, the first, qualitative part of the study would be used to focus the survey on the most relevant issues (mothers' feeding behaviours and reasons for these behaviours) and to help phrase the questions in an optimal way in order to obtain the information that is needed.

The second, quantitative part of the study would be used to find out what proportion of the mothers follow various practices and the reasons for their behaviours and whether certain categories of children (e.g., the younger ones or children from specific socio-economic categories) are more at risk than others.

The third, qualitative part of the study would provide feedback on the major findings of the survey. Do the conclusions make sense to women in the study area? Have certain aspects been overlooked when interpreting the data? What remedial action is feasible to improve practices related to feeding sick children?

It is also common to collect qualitative and quantitative data in a single questionnaire. Researchers collecting both types of data have to take care that they:

- do not include too many open-ended questions in large-scale surveys, making data analysis more complicated; and
- do not use inappropriate statistical tests on quantitative data generated by small-scale studies.

The Qualitative and Quantitative data are not fundamentally different- Qualitative data consists of words which can be coded quantitatively and Quantitative data consists of numbers which are based on qualitative judgement.

Because of this interrelationship between Qualitative and Quantitative data it is very important for the researcher who is administering the interview schedule to have a comprehensive idea of the variables to be explored in the study & ways of accounting for them and different scaling techniques so that he/she can extract as much information as is needed for studying and explaining the problem selected for research.

Variables

In statistics, variables refer to measurable attributes, as these typically vary over time or between individuals. Variables can be continuous (taking values from a continuum) or discrete (taking values from a defined set). Temperature is a continuous variable, while

number of legs of an animal is a discrete variable. This concept of a variable is widely used in the natural, medical and social sciences.

Variables classified according to levels of attributes measured:

(1)Categorical variables:
Variables that depict attributes or categories of a concept that cannot be reduced to a number or numerical scale; they vary in kind. There are two kinds of categorical variables:

- *Nominal variables* do not vary according to a specific order. The categories of nominal variables are simply names. Eg : Political party classification

- *Ordinal variables* vary according to a specific order, but the degrees of separation between their ranks cannot be numerically specified. Eg: Socio-economic status

(2)Numerical variables:
Variables that depict attributes or categories of a concept that can be reduced to a number or numerical scale; they vary in degree. There are two kinds of numerical variables:

- *Interval variables* vary according to a specific order; the degrees of separation between their ranks can be numerically specified, but no true zero point orders their measured variation. Eg: Intelligence

- *Ratio variables* vary according to a specific order; the degrees of separation between their ranks can be numerically specified, and a true zero point orders their measured variation. Eg: weight, height

Variables classified according to causal models:

- **Independent and Dependent variables**

In causal models, a distinction is made between **"Independent variables"** and **"Dependent variables"**, the latter being expected to vary in value in response to changes in the former. In other words, an independent variable is presumed to potentially affect a dependent one. In experiments, independent variables include factors that can be altered or chosen by the researcher independent of other factors.

For example, in an experiment to test whether the boiling point of water changes with altitude, the altitude is under direct control and is the independent variable, and the boiling point is presumed to depend upon it and is therefore the dependent variable.

Testing for the relationship of Independent and Dependent variables:

Selection of Statistical tests for studying the relationship of Independent and Dependent variable will depend upon (1) the number of independent and dependent variables

involved in the study (2) the nature of the Independent and Dependent variables- whether metric or non metric.
Given below are some examples:

Canonical correlation:
With canonical analysis the objective is to correlate *simultaneously several metric dependent variables and several metric independent variables*. The underlying principle is to develop a linear combination of each set of variables (both independent and dependent) to maximize the correlation between the two sets (Hair et al., 1995).

LOGIT:
Logit analysis is a special form of regression in which the *criterion variable is a non-metric, dichotomous (binary) variable*. While differences exist in some aspects, the general manner of interpretation is quite similar to linear regression (Hair et al., 1995).

Multiple Regression:
Multiple regression is the appropriate method of analysis when the research problem involves a *single metric dependent variable presumed to be related to one or more metric independent variables*. The objective of multiple regression analysis is to predict the changes in the dependent variable in response to the changes in the several independent variables (Hair et al., 1995).

Multiple Discriminant Analysis:

If the *single dependent variable is dichotomous (e.g., male-female) or multichotomous (e.g., high-medium-low) and therefore non-metric*, the multivariate technique of multiple discriminant analysis (MDA) is appropriate. As with multiple regression, the independent variables are assumed to be metric (Hair et al., 1995).

Multivariate Analysis of Variance (MANOVA):

Multivariate analysis of variance (MANOVA) is a statistical technique that can be used to simultaneously explore the *relationship between several categorical independent variables (usually referred to as treatments) and two or more metric dependent variables*. As such, it represents an extension of univariate analysis of variance (ANOVA). MANOVA is useful when the researcher designs an experimental situation (manipulation of several non-metric treatment variables) to test hypotheses concerning the variance in group responses on two or more metric dependent variables (Hair et al., 1995).

- **Intervening Variable**

An **Intervening variable** is a hypothetical concept that attempts to explain relationships between variables, and especially the relationships between independent variables and dependent variables. Intervening variables are not real things. They are interpretations of observed facts, not facts themselves. But they create the illusion of being facts.

Example: learning, memory, motivation, attitude, personality, traits, knowledge, understanding, thinking, expectation, intelligence, intention.

It is often distinguished from a hypothetical construct in that it has no properties other than those observed in empirical research. That is, it is simply a summary of the relationships observed between independent and dependent variables. A **Mediator variable (or mediating variable)** in statistics is a variable that describes how rather than when effects will occur by accounting for the relationship between the independent and dependent variables. A mediating relationship is one in which the path relating A to C is mediated by a third variable (B).

A mediating variable explains the actual relationship between the Independent and Dependent variables.

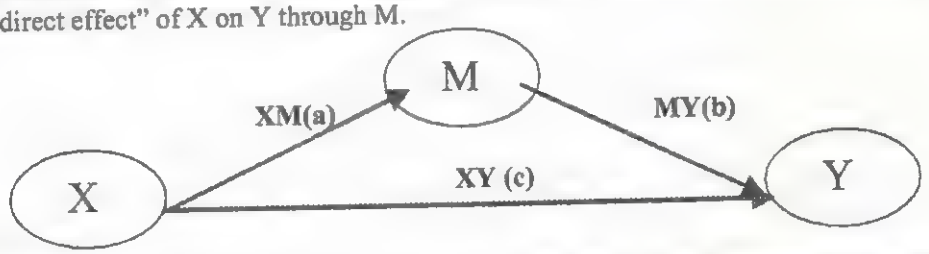
Let's look at the experiment by Tolman & Honzik (1930) on latent learning in rats, specifically, the group that received a reward every time they reached the goal box.

One of the Independent Variables was the number of practice trials the rats received. They got 1 trial per day, so each rat got an increasing number of trials.

The Dependent Variable was the number of wrong turns (errors) that rats made on a trial.

Latent learning is the mediating variable which explains the effect of practice trial on errors the rat made on trials.

Consider a model that proposes that some independent variable (X) is correlated with some dependent variable (Y) not because it exerts some direct effect upon the dependent variable, but because it causes changes in an intervening or mediating variable (M), and then the mediating variable causes changes in the dependent variable. Psychologists tend to refer to the X → M → Y relationship as "mediation." Sociologists tend to speak of the "indirect effect" of X on Y through M.



Baron and Kenny (1986) provide a clear explication of the meaning of mediating variables.

A variable functions as a mediator when it meets the following conditions: (a) variations in levels of the independent variable significantly account for variations in the presumed mediator (i.e., Path a), (b) variations in the mediator significantly account for variatons in the dependent variable (i.e., Path b), and (c) when Paths a and b are controlled, a previously significant relation between the independent and dependent variables is no

longer significant, with the strongest demonstration of mediation occurring when Path c is reduced to zero. If the residual Path c is not zero, this indicates the operation of multiple mediating factors. Because most areas of psychology, including social, treat phenomena that have multiple causes, a more realistic goal may be to seek mediators that significantly decrease Path c rather than eliminating the relations between the independent and dependent variables altogether. From a theoretical perspective, a significant reduction demonstrates that a given mediator is indeed potent, albeit not both a necessary and sufficient condition for an effect to occur.

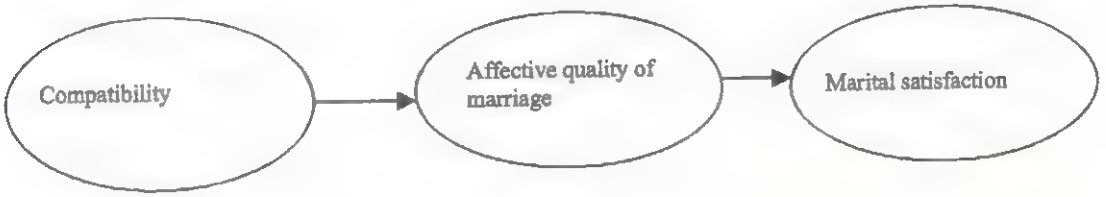
Mediating variables are often contrasted with **Moderating variables**, which pinpoint the conditions under which an independent variable exerts its effects on a dependent variable. A moderating relationship can be thought of as an interaction. It occurs when the relationship between variables A and B depends on the level of C.

Path Diagrams

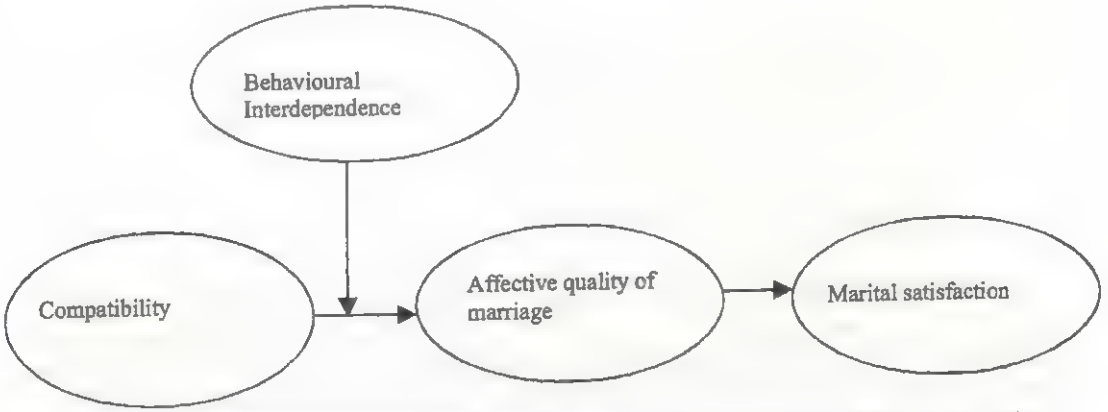
Path diagrams provide an easy and convenient way to represent linkages between and among constructs (cf., Loehlin, 1987).Path diagrams distinguish theoretical constructs from measured variables. Theoretical constructs are abstract by nature; they are often referred to as latent variables because they are not measured directly. "Measured variables" writes Falk (1987), "are actual observations and are frequently called markers, indicants, or manifest variables" (p. 14). Markers of compatibility might be similarity in gender-role attitudes, the extent to which partners like the same leisure activities, and their level of agreement in regard to religious matters. Theoretical constructs are represented in path diagrams as spheres, measured variables by squares. The distinction between theoretical constructs and measured variables is useful to keep in mind even when only one measure for each theoretical construct is used. The failure to find an empirical association between two measured variables, each standing for a different theoretical construct, does not necessarily mean the two theoretical constructs are unrelated. Other more valid and more reliable measures may demonstrate the hypothesized relationship. A path diagram can be used to show the linkages between theoretical constructs and between such constructs and measured variables. The connections can be shown using two kinds of arrows: (a) straight one-headed arrows representing unidirectional, or causal, relationships between variables; and (b) curved, two-headed arrows depicting covariation, or correlation. Straight arrows going both directions between variables can be used to show mutual influence. Numbered subscripts on variables may designate time periods (when the measures pertain to data gathered at different points in time). The first step in creating a path diagram, regardless of whether multiple measures are used, involves the creation of a diagram showing the linkages between the latent, or theoretical, variables. Such a diagram is called a Latent Variable Path Model, or the "inner model" (Falk, 1987).

The preliminary model shown below portrays a rather common formulation suggesting that compatibility affects marital satisfaction. The arrow drawn from compatibility to marital satisfaction suggests that compatibility has a causal effect on satisfaction. It is incumbent upon the theoretician to create a plausible rationale for linking the variables.

Compatibility might be posited to account, at least in part, for the affective quality of the marriage. The affective quality, in turn, may be hypothesized to affect satisfaction. This formulation suggests that the impact of compatibility on marital satisfaction is mediated through marital interaction. This set of propositions puts together ideas concerning the causes of marital satisfaction based upon compatibility theories and social learning theory.



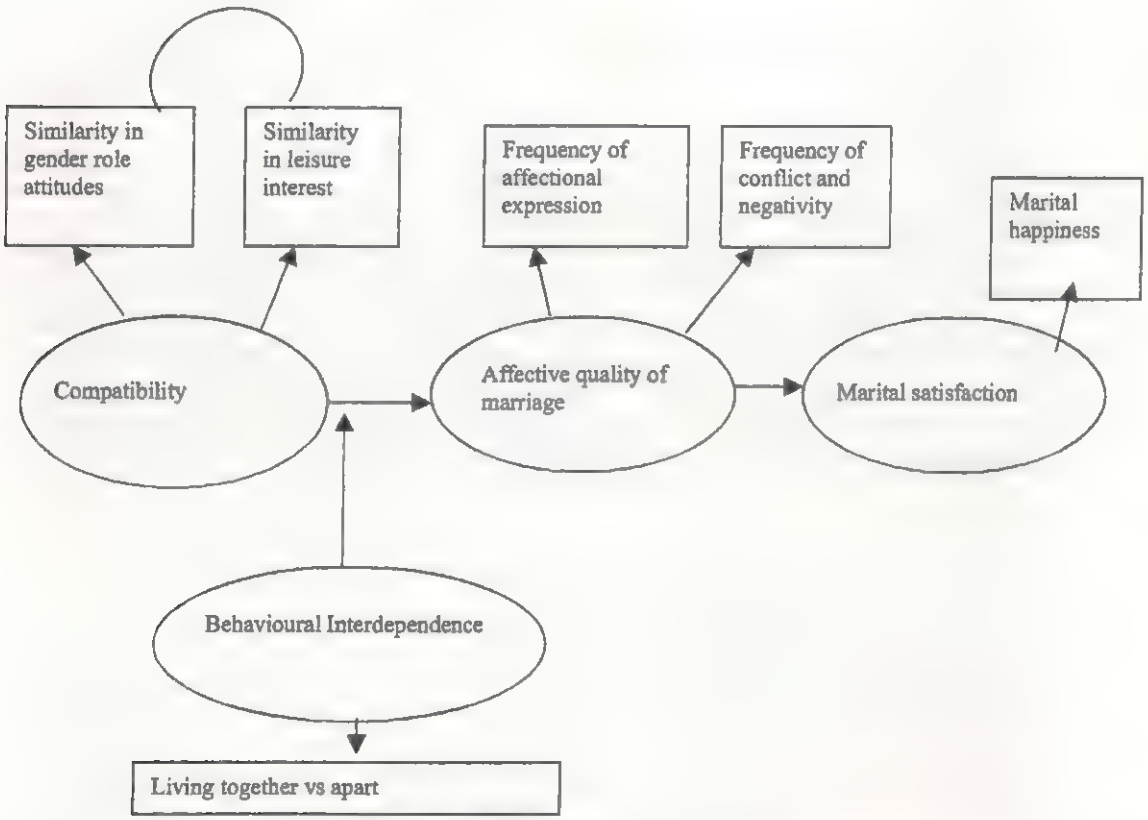
The next figure shows an even more elaborated formulation of the connection between compatibility and marital satisfaction. The line drawn from behavioral interdependence to the line connecting compatibility and affective quality of the marriage suggests that the relationship between the latter two constructs depends upon the extent of behavioral interdependence. In this case, it is suggested that compatibility has little or no connection with the affective quality of interaction when interdependence is low, but a strong connection when interdependence is high. Behavioral interdependence thus moderates the connection between compatibility and satisfaction.



The constructs should be ordered in the path diagram from left to right, with those at the left causally prior to, or predictive of, those on the right. The theoretical construct(s) to

the far left are generally thought of as exogenous (outside) because the model takes their values as "given," rather than as something to be explained. Compatibility is an exogenous variable. The variables that are shown to be influenced, either directly or indirectly, by exogenous variables are said to be endogenous. It should be clear that the causal chain could be extended farther back; another model, for example, might explain "compatibility" in terms of dating experience, the idea being that those who shop around will be more likely to select a more compatible partner.

The next step in constructing a path diagram involves showing the connection between the latent variables and the markers, or measured variables. It is frequently the case that each theoretical construct is measured with a single indicator. For purposes of illustration, however, we will develop a more complex model in the example shown below.



Two variables measuring compatibility are included - similarity in gender role attitudes and similarity in leisure interests. The curved two-headed arrow drawn between these two measures of compatibility portrays them as correlated, but not causally related. The rationale behind this depiction, as well as each of the other arrows in the model, needs to be articulated. It might be argued, for example, that couples who are "selected" with regard to gender role attitudes might also be selected in terms of leisure interests. If some

people more than others select mates on the basis of compatibility, then it would not be surprising to find that couples who are compatible in one regard would also be compatible in others.

The arrows drawn to similarity in gender role attitudes and leisure interests identify them as indicants of compatibility. The frequency of affectional expression and the frequency of conflict and negativity are shown as indicants of the affective quality of marriage. You may have noticed that affectional expression and negativity are not shown as correlated. The absence of the arrow in this instance is based on a review of previous research that shows no correlation between affectional expression and conflict/negativity among couples who are happy with their marriage (note that the PAIR Project partners are generally happy in the early years of marriage. If you were using data from the follow-up, however, you might carefully examine this correlation or lack thereof, since many partners are unhappy in the fourth phase). The spouses' living arrangement is used as an indicant of behavioral interdependence, suggesting that the connection between compatibility and the affective quality of marriage depends upon whether the spouses live together or have a commuting marriage. The right hand side of the model shows a single measure of satisfaction.

Testing for Mediating variables

MacKinnon, Lockwood, Hoffman, West, and Sheets (A comparison of methods to test mediation and other intervening variable effects, *Psychological Methods*, 2002, *7*, 83-104) reviewed 14 different methods that have been proposed for testing models that include intervening variables. Some of these are:

(1)Causal Steps. This is the approach that has most directly descended from the work of Judd, Baron, and Kenny and which has most often been employed by psychologists. Using this approach, the criteria for establishing mediation, which are nicely summarized by David Howell (*Statistical Methods for Psychology*, 6th ed., page 528) are:

- X must be correlated with Y.
- X must be correlated with M.
- M must be correlated with Y, holding constant any direct effect of X on Y.
- When the effect of M on Y is removed, X is no longer correlated with Y (complete mediation) or the correlation between X and Y is reduced (partial mediation).

Each of these four criteria are tested separately in the causal steps method:

- First you demonstrate that the zero-order correlation between X and Y (ignoring M) is significant.
- Next you demonstrate that the zero-order correlation between X and M (ignoring Y) is significant.
- Now you conduct a multiple regression analysis, predicting Y from X and M. The partial effect of M (controlling for X) must be significant.
- Finally, you look at the direct effect of X on Y. This is the Beta weight for X in the multiple regression just mentioned. For complete mediation, this Beta must

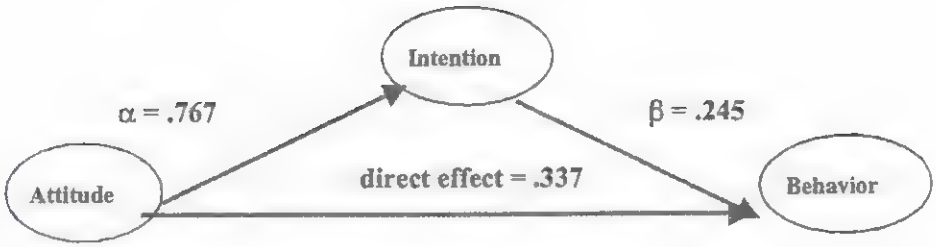
be (not significantly different from) 0. For partial mediation, this Beta must be less than the zero-order correlation of X and Y.

(2)Difference in Coefficients. These methods involve comparing two regression or correlation coefficients -- that for the relationship between X and Y ignoring M and that for the relationship between X and Y after removing the effect of M on Y. MacKinnon et al. describe a variety of problems with these methods, including unreasonable assumptions and null hypotheses that can lead one to conclude that mediation is taking place even when there is absolutely no correlation between M and Y.

(3) Product of Coefficients. One can compute a coefficient for the "indirect effect" of X on Y through M by multiplying the coefficient for path XM by the coefficient for path MY. The coefficient for path XM is the zero-order *r* between X and M. The coefficient for path MY is the Beta weight for M from the multiple regression predicting Y from X and M (alternatively one can use unstandardized coefficients).

One can test the null hypothesis that the indirect effect coefficient is zero in the population from which the sample data were randomly drawn. The test statistic (*TS*) is computed by dividing the indirect effect coefficient by its standard error, that is, $TS = \frac{\alpha\beta}{\sigma_{\alpha\beta}}$. This test statistic is usually evaluated by comparing it to the standard normal distribution. The most commonly employed standard error is **Sobel's (1982) first-order approximation**, which is computed as $\sqrt{\alpha^2\sigma_\beta^2 + \beta^2\sigma_\alpha^2}$, where α is the zero-order correlation or unstandardized regression coefficient for predicting M from X, $\sigma_\alpha{}^2$ is the standard error for that coefficient, β is the standardized or unstandardized partial regression coefficient for predicting Y from M controlling for X, and $\sigma_\beta{}^2$ is the standard error for that coefficient. An alternative standard error is **Aroian's (1944) second-order exact solution**, $\sqrt{\alpha^2\sigma_\beta^2 + \beta^2\sigma_\alpha^2 + \sigma_\alpha^2\sigma_\beta^2}$. Another alternative is **Goodman's (1960) unbiased solution**, in which the rightmost addition sign becomes a subtraction sign: $\sqrt{\alpha^2\sigma_\beta^2 + \beta^2\sigma_\alpha^2 - \sigma_\alpha^2\sigma_\beta^2}$.

MacKinnon et al. gave some examples of hypotheses and models that include intervening variables. One was that of Ajzen & Fishbein (1980), in which intentions are hypothesized to intervene between attitudes and behavior. Ingram, Cope, Harju, and Wuensch (Applying to graduate school: A test of the theory of planned behavior. *Journal of Social Behavior and Personality*, 2000, *15*, 215-226) tested a model which included three "independent" variables (attitude, subjective norms, and perceived behavior control), one mediator (intention), and one "dependent" variable (behavior). The model is simplified here dropping subjective norms and perceived behavioral control as independent variables. Accordingly, the mediation model (with standardized path coefficients) is:



Let us first consider the causal steps approach:

- Attitude is significantly correlated with behavior, $r = .525$.
- Attitude is significantly correlated with intention, $r = .767$.
- The partial effect of intention on behavior, holding attitude constant, falls short of statistical significance, $\beta = .245$, $p = .16$.
- The direct effect of attitude on behavior (removing the effect of intention) also falls short of statistical significance, $\beta = .337$, $p = .056$.

The causal steps approach does not, here, provide strong evidence of mediation, given lack of significance of the partial effect of intention on behavior. If sample size were greater, however, that critical effect would, of course, be statistically significant.

Now the Sobel/Aroian/Goodman tests is calculated. The statistics needed are the following:

- The zero-order unstandardized regression coefficient for predicting the mediator (intention) from the independent variable (attitude). That coefficient = .423.
- The standard error for that coefficient = .046.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.390	1.519		2.231	.030
	attitude	.423	.046	.767	9.108	.000

a. Dependent Variable: intent

- The partial, unstandardized regression coefficient for predicting the dependent variable (behavior) from the mediator (intention) holding constant the independent variable (attitude). That regression coefficient = 1.065.
- The standard error for that coefficient = .751.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.075	9.056		.008	.993
	attitude	.807	.414	.337	1.950	.056
	intent	1.065	.751	.245	1.418	.162

a. Dependent Variable: behav

For Aroian's second-order exact solution,

$$TS = \frac{\alpha\beta}{\sqrt{\alpha^2\sigma_\beta^2 + \beta^2\sigma_\alpha^2 + \sigma_\alpha^2\sigma_\beta^2}} = \frac{.423(1.065)}{\sqrt{.423^2(.751)^2 + 1.065^2(.046)^2 + .046^2(.751)^2}} = 1.3935$$

(4) Mackinnon et al. (1998) Distribution of Products $\frac{\alpha\beta}{\sigma_{\alpha\beta}}$. With this approach, one starts by converting both critical paths (α and β in the figure above) into *z* scores by dividing their unstandardized regression coefficients by the standard errors (these are, in fact, the *t* scores reported in typical computer output for testing those paths). For our data, that yields $Z_\alpha Z_\beta = 9.108 * 1.418 = 12.915$. For a .05 nondirectional test, the critical value for this test statistic is 2.18. Again, our evidence of mediation is significant.

(5) Bootstrap Analysis. Partick Shrout and Niall Bolger published an article, "Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations," in the *Psychological Bulletin* (2002, *7*, 422-445), in which they recommend that one use bootstrap methods to obtain better power, especially when sample sizes are not large.

Testing for moderating variables

Moderated regression model discussed by Zedeck (1971) test the effect of moderating variables on the relationship of X (IV) and Y (DV)
The moderated regression equation is:
Y= bo +bi Xi + bj MO j + bij Xi MOj
Where
Y= Dependent variable score
Xi= Independent variable score
MOj= Moderating variable score
Xi MOj = Independent variable-Moderating variable score interaction,

In each regression Xi, MOj and Xi – MOj were entered hierarchically in that order. According to Zedeck (1971) moderator effect is present when the independent predictor model and the moderated regression model both differ significantly from the zero order correlation and additionally are significantly different from one another.

Scaling Techniques

In the social sciences, scaling is the process of measuring or ordering entities with respect to quantitative attributes or traits. For example, a scaling technique might involve estimating individuals' levels of extraversion, or the perceived quality of products. Certain methods of scaling permit estimation of magnitudes on a continuum, while other methods provide only for relative ordering of the entities.

Data types

The type of information collected can influence scale construction. Different types of information are measured in different ways.

Some data are measured at the **nominal level**. That is, any numbers used are mere labels: they express no mathematical properties. Examples are SKU inventory codes and UPC bar codes. In nominal group members of any two groups are never equivalent but all members of any one group are always equivalent. In case of nominal measurement admissible statistical operations are counting or frequency, percentage, proportion, mode and coefficient contingency. Addition, subtraction, multiplication and division are not possible.

Some data are measured at the **ordinal level**. Numbers indicate the relative position of items, but not the magnitude of difference. An example is a preference ranking. The permissible statistical operation in ordinal measurement are median, percentile and rank correlation coefficient plus all those which are permissible for nominal measurement.

Some data are measured at the **interval level**. Numbers indicate the magnitude of difference between items, but there is no absolute zero point. Examples are attitude scales and opinion scales. The common statistics used in such measurement are arithmetic mean, standard deviation, Pearson r and other statistics based upon them. t-tests and f tests are also applied.

Some data are measured at the **ratio level**. Numbers indicate magnitude of difference and there is a fixed zero point. Ratios can be calculated. Examples include: age, income, price, costs, sales revenue, sales volume, and market share. All statistics including coefficient of variation can be utilized.

Scale construction decisions

- What level of data is involved (nominal, ordinal, interval, or ratio)?
- What will the results be used for?
- What types of statistical analysis would be useful?
- Should you use a comparative scale or a noncomparative scale?
- How many scale divisions or categories should be used (1 to 10; 1 to 7; -3 to +3)?
- Should there be an odd or even number of divisions? (Odd gives neutral center value; even forces respondents to take a non-neutral position.)
- What should the nature and descriptiveness of the scale labels be?
- What should the physical form or layout of the scale be? (graphic, simple linear, vertical, horizontal)
- Should a response be forced or be left optional?

Comparative and noncomparative scaling

With comparative scaling, the items are directly compared with each other (example : Do you prefer Pepsi or Coke?). In noncomparative scaling each item is scaled independently of the others (example : How do you feel about Coke?).

Comparative scaling techniques

- **Pairwise comparison scale** - a respondent is presented with two items at a time and asked to select one (example : Do you prefer Pepsi or Coke?). *This is an ordinal level technique when a measurment model is not applied.* Krus and Kennedy (1977) elaborated the paired comparison scaling within their domain-referenced model.

- **Rank-order scale** - a respondent is presented with several items simultaneously and asked to rank them (example : Rate the following advertisements from 1 to 10.). *This is an ordinal level technique.*

- **Constant sum scale** - a respondent is given a constant sum of money, script, credits, or points and asked to allocate these to various items (example : If you had 100 Yen to spend on food products, how much would you spend on product A, on product B, on product C, etc.). *This is an ordinal level technique.*

- **Bogardas social distance scale** - measures the degree to which a person is willing to associate with a class or type of people. It asks how willing the respondent is to make various associations. The results are reduced to a single score on a scale. There are also non-comparative versions of this scale.

- **Q-Sort scale** - Up to 140 items are sorted into groups based a rank-order procedure.

- **Guttman scale** - This is a procedure to determine whether a set of items can be rank-ordered on an unidimensional scale. It utilizes the intensity structure among several indicators of a given variable. Statements are listed in order of importance. The rating is scaled by summing all responses until the first negative response in the list.

Non-comparative scaling techniques

- **Continuous rating scale (also called the graphic rating scale)** - respondents rate items by placing a mark on a line. The line is usually labeled at each end. There are sometimes a series of numbers, called scale points, (say, from zero to 100) under the line. Scoring and codification is difficult.

- **Likert scale** - Respondents are asked to indicate the amount of agreement or disagreement (from strongly agree to strongly disagree) on a five- or seven-point scale. The same format is used for multiple questions.

- **Phrase completion scales** - Respondents are asked to complete a phrase on an 11-point response scale in which 0 represents the absence of the theoretical construct and

10 represents the theorized maximum amount of the construct being measured. The same basic format is used for multiple questions.

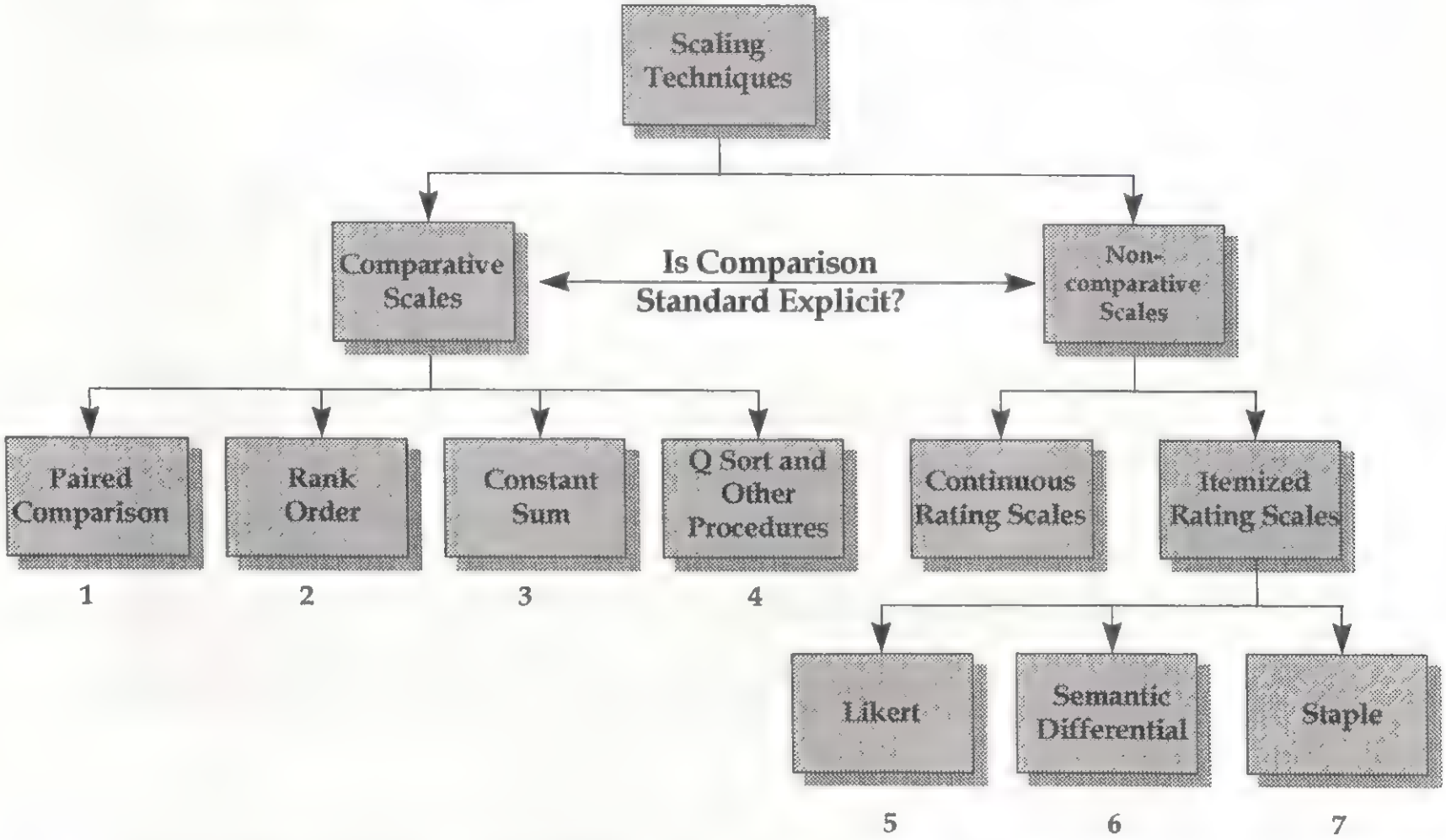
- **Semantic differential scale** - Respondents are asked to rate on a 7 point scale an item on various attributes. Each attribute requires a scale with bipolar terminal labels.

- **Stapel scale** - This is a unipolar ten-point rating scale. It ranges from +5 to -5 and has no neutral zero point.

- **Thurstone scale** - This is a scaling technique that incorporates the intensity structure among indicators.

Some examples of the different scaling techniques are as follows:

Figure 10.1 A Classification of Scaling Techniques



NOTE: Above scales provide different types of data (i.e., nominal, ordinal, etc.)

TABLE 11.4 **A Paired Comparison Scale for Suntan Products**

14. Thinking about sun products in general, here are some characteristics used to describe them. Please tell me which characteristic in each pair is more important to you when selecting a sun care product.

a. Tans evenly	b. Tans without burning
a. Prevents burning	b. Protects against burning and tanning
a. Good value for the money	b. Goes on evenly
a. Not greasy	b. Does not stain clothing
a. Tans without burning	b. Prevents burning
a. Protects against burning and tanning	b. Good value for the money
a. Goes on evenly	b. Tans evenly
a. Prevents burning	b. Not greasy

Figure 10.2 Obtaining Shampoo Preferences Using Paired Comparisons

Instructions

We are going to present you with 10 pairs of shampoo brands. For each pair, please indicate which one of the two brands of shampoo in the pair you would prefer for personal use.

Recording Form

	Jhirmack	Finesse	Vidal Sassoon	Head & Shoulders	Pert
Jhirmack		0	0	1	0
Finesse	1[a]		0	1	0
Vidal Sassoon	1	1		1	1
Head & Shoulders	0	0	0		0
Pert	1	1	0	1	
Number of Times Preferred[b]	3	2	0	4	1

[a] A 1 in a particular box means that the brand in that column was preferred over the brand in the corresponding row. A 0 means that the row brand was preferred over the column brand.

[b] The number of times a brand was preferred is obtained by summing the 1's in each column.

Limit on number of paired comparisons

Concept of "Intransitive paired comparisons"

A > B

B > C

C > A (should not be possible)

Ordinal Interval (Assumed)
Paired Comparison Scale

PAIRED-COMPARISON DATA FOR 5 BRANDS OF CAKE MIX

Matrix A

	A	B	C	D	E
A	—	.90	.64	.14	.27
B	.10	—	.32	.02	.21
C	.36	.68	—	.15	.36
D	.86	.98	.85	—	.52
E	.73	.79	.64	.48	—

Matrix B

	A	B	C	D	E
A	—	1	1	0	0
B	0	—	0	0	0
C	0	1	—	0	0
D	1	1	1	—	1
E	1	1	1	0	—
Total	2	4	3	0	1

Entries represent proportion of respondents preferring "column" brand to "row" brand (i.e., 90 percent prefer Brand B to Brand A)

Figure 10.3 Preference for Toothpaste Brands Using Rank Order Selling

Instructions

Rank the various brands of toothpaste in order of preference. Begin by picking out the one brand that you like most and assign it a number 1. Then find the second most preferred brand and assign it a number 2. Continue this procedure until you have ranked all the brands of toothpaste in order of preference. The least preferred brand should be assigned a rank of 10.

No two brands should receive the same rank number.

The criterion of preference is entirely up to you. There is no right or wrong answer. Just try to be consistent.

Brand	Rank Order
1. Crest	3
2. Colgate	
3. Aim	1
4. Gleem	
5. Macleans	2
6. Ultra Brite	
7. Close Up	4
8. Pepsodent	
9. Plus White	
10. Stripe	

Natural Technique in Marketing?

Limit on number of brands that can be ranked

A Series of Rank-Order Scales Used to Evaluate Eye Shadows

Please rank the following eye shadows with 1 being the brand that best meets the characteristic being evaluated and 6 the worst brand on the characteristic being evaluated. Let's begin with the idea of having high-quality compacts or containers. Which brands would rank as having the highest quality compacts or containers? Which is second? (RECORD BELOW.)

	Having High-Quality Container	Having High-Quality Applicator	Having High-Quality Eye Shadow
Avon			
Cover Girl	1	3	1
Estee Lauder		1	2
Maybelline	2	2	3
Natural Wonder			
Revlon	3		

– Possible to include a phony brand
– Create "composite" rank order

Figure 10.4 Importance of Toilet Soap Attributes Using a Constant Sum Scale

Instructions

Below are eight attributes of toilet soaps. Please allocate 100 points among the attributes so that your allocation reflects the relative importance you attach to each attribute. The more points an attribute receives, the more important the attribute is. If an attribute is not at all important, assign it zero points. If an attribute is twice as important as some other attribute, it should receive twice as many points.

Form

	AVERAGE RESPONSES OF THREE SEGMENTS		
Attribute	Segment I	Segment II	Segment III
1. Mildness	8	2	4
2. Lather	2	4	17
3. Shrinkage	3	9	7
4. Price	53	17	9
5. Fragrance	9	0	19
6. Packaging	7	5	9
7. Moisturizing	5	3	20
8. Cleaning Power	13	60	15
Sum	100	100	100

-- Difficult and unnatural?
-- Gives the best scale properties

TABLE 11.5 **A Constant Sum Scale Used in a Tennis Sportswear Study**

Below are seven characteristics of women's tennis sportswear. Please allocate 100 points among the characteristics in such a way that the allocation represents the importance of each characteristic to you. The more points that you assign to a characteristic, the more important it is. If the characteristic is totally unimportant, you should not allocate any points to it. When you've finished, please double check to make sure that your total adds to 100.

CHARACTERISTICS OF TENNIS SPORTSWEAR	NUMBER OF POINTS
Is comfortable to wear	_____
Is durable	_____
Is made by well-known brand or sports manufacturers	_____
Is made in the U.S.A.	_____
Has up-to-date styling	_____
Gives freedom of movement	_____
Is a good value for the money	_____
	_____
	100 points

Constant Sum Use with Paired-Comparison Scale

Ratio

Now, we would like you to divide 11 points, or "chips," between each pair of hand and body lotions listed below. You can divide the 11 chips any way you like, depending on how much more you like one product than the other. Some possible combinations are 11 & 0, 1 & 10, 9 & 2, 3 & 8, 7 & 4, or 5 & 6. *The two numbers you assign to the products in each pair must add up to eleven.* Always give the product you like more in each pair the larger of the two numbers. Now, please rate each pair of products going *across* the page. The number of chips you decide to give a product should be recorded in the box next to it.

		Chips			Chips
A.	Agree	8	B.	Alberto VO-5	3
B.	Alberto VO-5	7	C.	Body on Tap	4
A.	Agree	9	C.	Body on Tap	2

— Bias created by order in which brands are presented?

— Difficult to do?

Q-Sort Procedure

Ordinal

The deck in front of you contains pictures of 75 magazines. Please choose the 9 magazines you most prefer of the 75. Once you have selected the 9 you most prefer, please list the magazine names on the form provided under the column "Prefer Most." Now, we would like you to select the 9 magazines you *least* prefer from the remaining 66 magazines. Please list these nine under the column "Prefer Least." Of the remaining 57 magazines, please select the 15 you most prefer and list these 15 under the column labeled "Like." Of the remaining 42 magazines, please choose the 15 you like the least and place these names under the column "Dislike." There should be 27 magazines remaining; please list the titles of these magazines under the column "Neutral."

Tabulation Sheet

Prefer Most	*Like*	*Neutral*	*Dislike*	*Prefer Least*
(9)	(15)	(27)	(15)	(9)

Non-Comparative Scales

ITEMIZED (Graphic) Ordinal

Scale A

ITEMIZED (Verbal) Ordinal

Scale B

- ☐ **Very favorable**
- ☐ **Somewhat favorable**
- ☐ **Indifferent**
- ☐ **Somewhat unfavorable**
- ☐ **Very unfavorable**

CONTINUOUS (Graphic) Ratio

Scale C

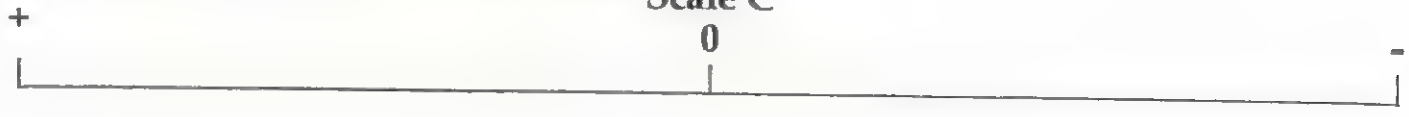


TABLE 11.1 Itemized Rating Scales Used in National Surveys

SCALE A

Now, I'd like to ask you about just two watches specifically. The first one is the SEARS watch. I'm going to mention some characteristics of watches, and, as I mention each one, please tell me whether you think the SEARS watch is (HAND RESPONDENT RATING CARD) excellent, very good, good, fair, or poor.

. . . for the particular characteristic?

The first characteristic is (READ CHARACTERISTIC CIRCLED BELOW). Do you feel that the SEARS watch is excellent, very good, good, fair, or poor for (CHARACTERISTIC)?

(CONTINUE FOR*ALL* CHARACTERISTICS BELOW)

STARTING POINT		EXCELLENT	VERY GOOD	GOOD	FAIR	POOR
X	Value for the money	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
X	Brand name	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
X	Accuracy	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
X	Durability	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
X	Manufacturer's reputation	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
X	After-sales service	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1
X	Styling	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1

SCALE B

9. Which statement on this card (HAND RESPONDENT CARD B) best describes the present condition of your hair?

1 ( ) Very damaged
2 ( ) Somewhat damaged
3 ( ) Slightly damaged
4 ( ) Not at all damaged

TABLE 11.1 **Itemized Rating Scales Used in National Surveys—*continued***

SCALE C

Now, I would like to get your opinion on Stridex Cleansing Pads on some characteristics. (HAND RATING CARD) Using the phrases on the card, please tell me which one best indicates how much you agree or disagree that Stridex Cleansing Pads . . . (START WITH CHECKED CHARACTERISTICS AND CONTINUE UNTIL ALL ARE ASKED)

START	AGREE STRONGLY	AGREE SOMEWHAT	DISAGREE SOMEWHAT	DISAGREE STRONGLY
( ) Help prevent blemishes	___9-4	___-3	___-2	___-1
( ) Help to clear up blemishes	___10-4	___-3	___-2	___-1
( ) Are convenient to use	___11-4	___-3	___-2	___-1
( ) Are not irritating	___12-4	___-3	___-2	___-1
( ) Leave face feeling fresh	___13-4	___-3	___-2	___-1
( ) Make you feel confident you are doing everything you can to help your skin look good	___14-4	___-3	___-2	___-1

SCALE B

Very Very Good **Very Very Poor**

SOURCE: Scale D is adapted from Fred Cutler, "To Meet Criticisms of TV Ads, Researchers Find New Ways to Measure Children's Attitudes," *Marketing News* (January 27, 1978), p.16, published by the American Marketing Association.

TABLE 11.2 **Selected Itemized Rating Scales**

PURCHASE INTENT				
Definitely will buy	Probably will buy	Probably will not buy		Definitely will not buy
LEVEL OF AGREEMENT				
Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
QUALITY				
Very good	Good	Neither good nor bad	Fair	Poor
DEPENDABILITY				
Completely dependable	Somewhat dependable	Not very dependable		Not dependable at all
STYLE				
Very stylish	Somewhat stylish	Not very stylish		Completely unstylish
SATISFACTION				
Completely satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Completely dissatisfied
COST				
Extremely expensive	Expensive	Neither expensive nor inexpensive	Slightly inexpensive	Very inexpensive
EASE OF USE				
Very easy to use	Somewhat easy to use	Not very easy to use		Difficult to use
COLOR BRIGHTNESS				
Extremely bright	Very bright	Somewhat bright	Slightly bright	Not bright at all
MODERNITY				
Very modern	Somewhat modern	Neither modern nor old-fashioned	Somewhat old-fashioned	Very old-fashioned

TABLE 11.6

A Likert Scale for Persons with Foot Odor Problems Who Have Not Tried Johnson's Odor-Eaters

(SHOW CARD J) Now, I would like to find out your impressions about Johnson's Odor-Eaters, which you said you were familiar with but had not tried. As I read each characteristic, please tell me, using the statements on this card, if you strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree.

	STRONGLY AGREE	AGREE	NEITHER AGREE NOR DISAGREE	DISAGREE	STRONGLY DISAGREE
They might make my feet feel hot	5	4	3	2	1
I am satisfied with what I am using	5	4	3	2	1
My problem is not serious enough	5	4	3	2	1
Too much trouble to cut them to fit to size	5	4	3	2	1
Price is too expensive	5	4	3	2	1
Might make my shoes too tight	5	4	3	2	1
I'm embarrassed to buy them	5	4	3	2	1
The advertising has not convinced me that the product is effective	5	4	3	2	1
Other insoles I've tried didn't work	5	4	3	2	1
Foot sprays work better	5	4	3	2	1
Foot powders work better	5	4	3	2	1
I've never used an insole	5	4	3	2	1
Wouldn't last more than a couple of weeks	5	4	3	2	1
Would look unattractive in my shoes	5	4	3	2	1
Would have to buy more than one pair	5	4	3	2	1
Would have to move them from one pair of shoes to another	5	4	3	2	1
No product for foot odor works completely	5	4	3	2	1
They might get too wet from perspiration	5	4	3	2	1
Don't know what an insole would feel like in my shoe	5	4	3	2	1

Card J

Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree

Likert Scale Items

ORDINAL

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
1. The commercial was soothing	5	4	3	2	1
2. The commercial was not entertaining	5	4	3	2	1
3. The commercial was insulting	(5)	4	3	2	1
4. The commercial was silly	5	(4)	3	2	1
5. The commercial was too "hard-sell"	5	4	3	2	1
6. The characters in the commercial were realistic	5	4	3	2	1
7. The commercial was not creative	5	4	3	2	1
8. The commercial clearly demonstrated the product's advantages	5	4	3	(2)	1
9. I will remember this commercial	5	(4)	3	2	1
10. The commercial had meaning to me personally	5	4	3	(2)	1

— Statements should be "balanced" (i.e., good & bad)
— Include "duplicate" statements

Semantic Differential Scale

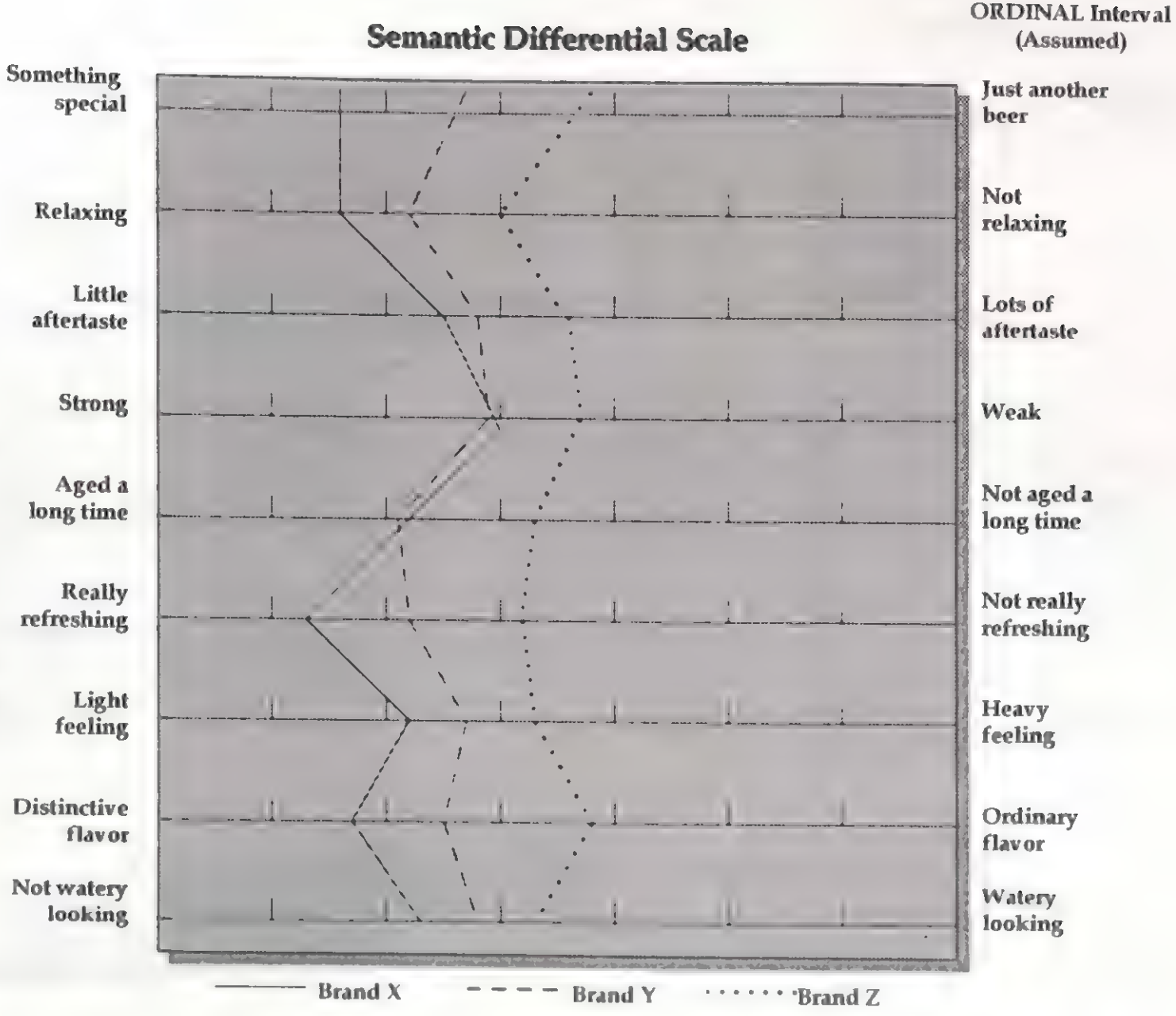
ORDINAL

In my opinion, the shaving experience using shaving cream code 4-11 was

Likable							X	Unlikable
Good								Bad
Cold				X				Hot
Unpleasant								Pleasant
Sharp								Dull
Useful								Useless
Boring			X					Interesting
Beneficial						X		Harmful
Attractive								Unattractive
Active								Passive

— Choice of adjectives?
— Used for image measurement
— Comparison with Staple scale?

Semantic Differential Scale



Example of a Staple Scale

Select a plus number for words that you think describe the bank accurately. The more accurately you think the word describes the company, the larger the plus number you should choose. Select a minus number for words you think do not describe the bank accurately. The less accurately you think the word describes the institution, the larger the minus number you should choose.

+5	+5
+4	+4
+3	+3
+2	+2
+1	+1
Friendly Personnel	Competitive Loan Rates
-1	-1
-2	-2
-3	-3
-4	-4
-5	-5

Staple Scale

ORDINAL Interval
(Assumed)

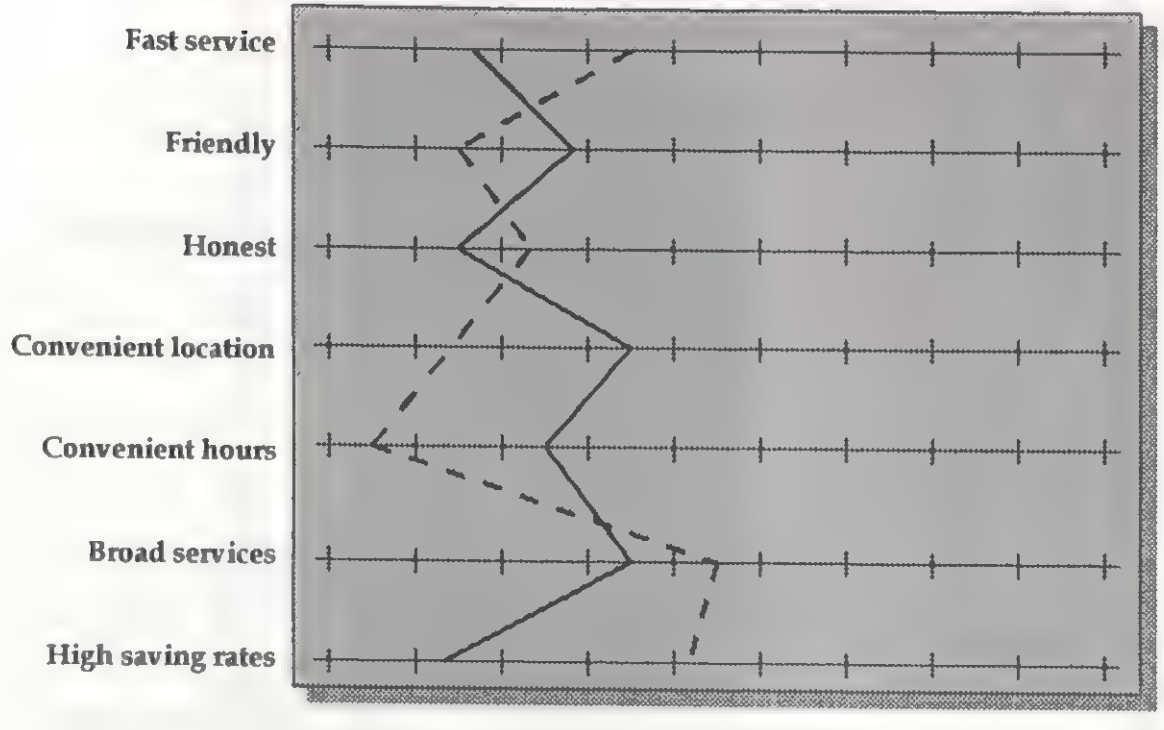


Figure 11.2 Rating Scale Configurations

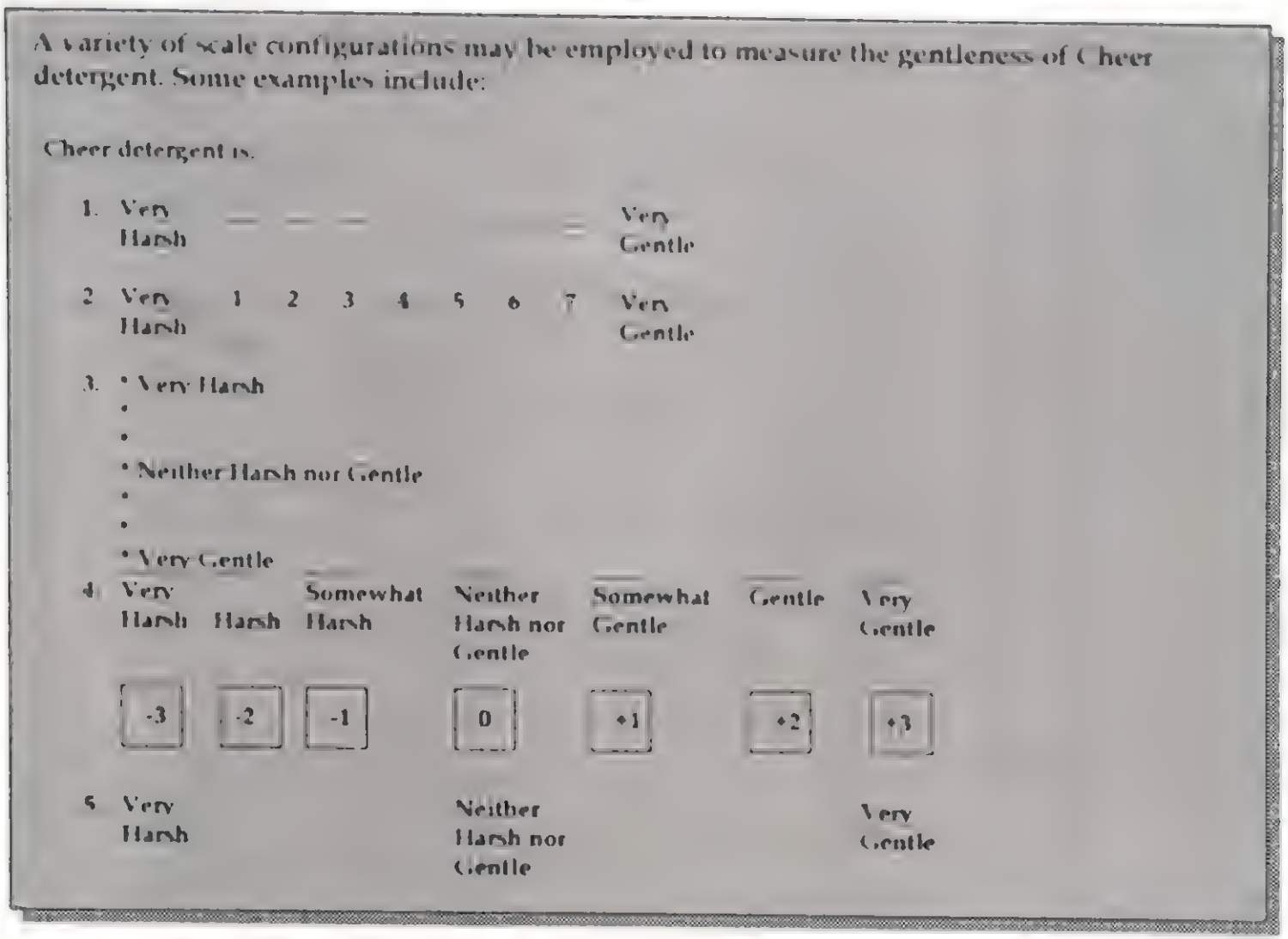


TABLE 11.7 Purchase Intent Scale and Related Questions for an In-Home Product Placement of Fly Traps

21. If a set of three traps sold for approximately $1.00 and was available in the stores where you normally shop, would you:

	(51)
definitely by the set of traps	1
probably buy	2
probably not buy--SKIP TO Q23	3
definitely not buy--SKIP TO Q23	4

22. Would you use the traps (a) instead of or (b) in addition to existing products?

	(52)
instead of	1
in addition to	2

23. Would you recommend this product to your friends?

	(53)
definitely	1
probably	2
probably not	3
definitely not	4

Dollar Metric Scale

With respect to fruit juice, which container do you prefer?				How much more in cents would you be willing to pay for the preferred container?
glass	✓	can		$.07
box		plastic	✓	.06
glass	✓	box		.07
plastic		glass	✓	.02
can	✓	box		.03
plastic		can	✓	.05

- **Used when price information is needed**
- **Can induce "bargaining behavior"**

Rating Scale Configurations

Different types of rating scales

1. Indicate your overall opinion about Safeway by placing a ☑ mark at an appropriate position on the line below:

   Very bad — Very good

   |——|

2. Indicate your overall opinion about Safeway by placing a ☑ mark in the category that best summarizes your feelings.

   Very bad — Very good

   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |

3. Indicate your overall opinion about Safeway by checking one of the following categories:

Very bad		Bad		Neither bad nor good		Good		Very good
[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]	[ ]
1	2	3	4	5	6	7	8	9

4. Which of the following best describes your overall opinion of Safeway?

Terrible	Poor	Fair	Good	Very good	Excellent
[ ]	[ ]	[ ]	[ ]	[ ]	[ ]

5. What is your overall rating of Safeway in comparison with other supermarkets in your area?

Much worse	Worse	About the same	Better	Much better
[ ]	[ ]	[ ]	[ ]	[ ]

6. Rank the following by placing a 1 beside the store you think is best overall, a 2 beside the store you think is second best, and so on:

   Kroger _____ Tom Thumb _____
   Piggly Wiggly _____ Winn-Dixie _____
   Safeway _____

7. In each of the following pairs, which store do you think is better? (Please check one store within each pair.)

   _____Kroger or _____Safeway
   _____Safeway or _____Piggly Wiggly
   _____Tom Thumb or _____Safeway
   _____Safeway or _____Winn-Dixie

8. Allocate a total of 100 points among the following stores, depending on how favorable you feel toward each; the more highly you think of each store, the more points you should allocate to it. (Please check that the allocated points add up to 100.)

   Kroger _____points
   Piggly Wiggly _____points
   Safeway _____points
   Tom Thumb _____points
   Winn-Dixie _____points

TABLE 11.1 Summary of Itemized Rating Scale Decisions

1. Number of categories	While there is no single, optimal number, traditional guidelines suggest that there should be between five and nine categories.
2. Balanced vs. unbalanced	In general, the scale should be balanced to obtain objective data.
3. Odd or even number of categories	If a neutral or indifferent scale response is possible for at least some of the respondents, an odd number of categories should be used.
4. Forced versus nonforced	In situations where the respondents are expected to have no opinion, the accuracy of data may be improved by a nonforced scale.
5. Verbal description	An argument can be made for labeling all or many scale categories. The category descriptions should be located as close to the response categories as possible.
6. Physical form	A number of options should be tried and the best one selected.
7. Category numbering to accompany verbal description	
8. Reference point or comparison standard being used by the respondemt	

Scaling Judgements

1. The scaling of items to a questionnaire

Answers to the queries or statements in most questionnaires admit of several possible replies, such as Yes, No?; or Most, Many, Some, Few, No; or there are four or five answers one of which is to be checked. It is often desirable to weigh these different selections in the degree of divergence from the typical answer which they indicate. First it is assumed that the attitude expressed in answering a given proposition is normally distributed. From the percentage who accept each of the possible answers to a question or statement a δ equivalent is found out which expresses the value or weight to be given to that answer. In order to avoid negative values, each δ weight can be expressed as a δ distance from-3.00δ .One advantage of a δ scaling is that the units of the scale are equal and hence may be compared from item to item or from scale to scale. Moreover a δ scaling gives a more accurate picture of the extent to which extreme or biased opinion on a given question are divergent from a typical opinion.

2.Scaling ratings in terms of the normal curve

In many psychological problems individuals are judged for their possession of characteristics or attributes not readily measured by tests. Honesty, interest in one's work, tactfulness, originality are illustrations of such traits. Suppose Two judges Jude 1 and Judge 2 have rated a group of 40 employees for their 'honesty' on a 5 point rating scale where a rating of A means the trait is possessed in marked degree and rating of E means it is almost if not completely absent and ratings of B,C and D means intermediate degrees of honesty.Assume the percentage of employees assigned each rating is as shown below

Honesty					
Judges	A	B	C	D	E
Judge 1	10%	15%	50%	20%	5%
Judge 2	20%	40%	20%	10%	10%

It is obvious that second judge rates more leniently than Judge 1 and that rating of A by Judge 1 may not mean the same degree of honesty as a rating of A by Judge 2.It is possible to assign numerical values to these ratings so as to make them comparable from Judge to Judge by transforming it to δ equivalents provided it is assumed that honesty is normally distributed in the population and one Judge is as competent as the other.

3. Changing order of merit into numerical scores

It is often desirable to transmute orders of merit into units of amounts or scores.This may be done by means of tables if the assumption of the normality of the trait is justified.It is possible to assign each person holding a rank a score on a scale of 100 points by the formula is Percent position= 100(R-.5)/N
Where R=rank of individual in the series
And N=Number of individuals ranked.

Incomplete order of merit ratings can be combined to give the final order of merit with the above formula

Suppose the data is as follows

	Students					
	A	B	C	D	E	F
Judge1	1	2	3	4	5	6
Judge 2		2		1		3
Judge 3	2		1		3	4

All scores have been transmuted, separate scores may be combined and averaged to give the final order of merit

	Students					
	A	B	C	D	E	F
Judge1	1	2	3	4	5	6
score	77	63	54	46	37	23
Judge 2		2		1		3
score		50		69		31
Judge 3	2		1		3	4
score	56		73		44	27
Sum of scores	133	113	127	115	81	81
Mean	67	57	64	58	41	27
Order of Merit	1	4	2	3	5	6

This method is useful in the case of those attributes which are not easily measured by ordinary methods, but for which individuals may be arranged in order of merit.It is also valuable for correlation when the only available criterion for a given ability or aptitude is a set of ranks. Transmuted scores can be averaged or combined like other test scores.

A comprehensive understanding of all the above areas will help the researcher to design the proper schedule and utilize the information obtained from it for better interpretation of the selected research problem.

References

- Baron, R.M and Kenny, DA (1986)The moderator-mediator variable distinction in social psychological research:Conceptual, strategic and statistical consideration, Journal of Personality and Social psychology, Vol 51, No6, 1173-1182
- Emory, W. C (1980), Business Research Methods, Irwin,.
- Falk, R.F (1987), A primer for soft modeling, Institute of Human Development, University of California, Berkeley;

- Hair, J. H. Jr., Anderson, R. E., Tatham, R. L., Black, William C., (1995) Multivariate Data Analysis, Prentice Hall,.
- Kerlinger, F.N.(1986). Foundations of Behavioral Research, Harcourt Brace Jovanovich,
- Loehlin, J.C (1987), Latent variable models: An introduction to factor, path, and structural analysis. Hillsdale, NJ: Erlbaum.
- Zedeck, S (1971) Problems with the use of moderator variables, Psychological Bulletin, Vol 76,pp295-310.

A look into how ubiquitous the role of non – response is: Some preventive as well compensatory measures.

Pulakesh Maiti
Economics Research Unit
Indian Statistical Institute
Kolkata

Summary: The practice of collection of statistics in India for commercial or government needs is age old. There are historical evidence of surveys through a rudimentary statistical system in India during the Hindu-Budhist period followed by a more mature system when the Moghuls ruled India. A rapid growth took place during the British period. Further growth and modernization with focus on Country's socio-economic progress occurred after India became independent in 1947.

In the Arthasastra by Kautilya (321 – 396 B.C.), which literally means a treatise on economics, one gets an account of data collection,

"It is duty of Gopa, village accountant, to attend the account of five or ten villages as ordered by the collector general..... Also, having numbered the houses as tax paying or non-tax paying, he shall not only register the total number of inhabitants of all the four castes in each village, but also keep an account of the exact number of cultivators, cowherds, merchants, artisans, labourers, slaves and biped and quadrupped animals, fixing the same time the amount of gold, free labour, toll and fines, that can be collected for it (each household)". We have also evidence in the great Indian Epic, the Mahabharata (Mahalanobis (1954), Halden (1957), Godambe (1976), Ghosh, Maiti (1999), Chatterjee (2003)]. A similar kind of picture on the detailed method of data collection is obtained through Ain-I-Akbari, a treatise on land revenue which was written by Abul Fazal, a famous scholar and humanist during the period of Akbar, the great Moghul emperor (1556 – 1605 A.D.).

No doubt such a practice was in force not only in India, but in many other countries especially in European Countries. Acceptance of sample surveys for data collection came into being in some countries in the last quarter of the 19th century, and since then, the history of studies of surveys is characterized by increasing use of surveys, the development of probability sampling methods, and a growing awareness about the nature and extent of errors that can affect the survey results.

Thus, in developing sampling design and making analysis of complex surveys, one should not focus narrowly on sampling and sampling errors, but rather he should take a broader view of the total survey process. We shall see that sampling errors are only one source of error and that indeed with large samples, they are always less important. One should therefore think of total error models for surveys, for non-coverage, non-response, response biases and variances and measurement errors of other sources.

1. Introduction

The data are neither error free not desirable at a given level of effort, an effort being measured in terms of time and cost. Thus the data are always contaminated, and hence it is an ethical necessity to provide tolerable margin of errors associated with the estimates to the users of data and the history of studies of surveys is therefore characterised by the increasing use of surveys, the development of probability sampling methods, and a growing awareness about the nature and extent of errors and an increasing sophistication of methods used for gathering information and controlling of errors. Surveys differ by the

purpose, subject matter and the type of methods used in data gathering operation [Moser and Kalton, (1972)]. Usually a distinction was made between scientific bodies based on observation data and those based on experimental data [Murthy, (1967), Jessen, (1978)]. To social scientists, surveys are restricted to a study of human populations [Babbie, (1973), Warwick and Lininger (1975), Hoinville etal (1978), Backstorm and Hursh ce'sar (1981)], while others recognized the role of surveys in studying institutions, agricultural or industrial production, business, inventions and so on [Hansen etal, (1953), Deming (1960), Dalenius (1974), Murthy (1963)].

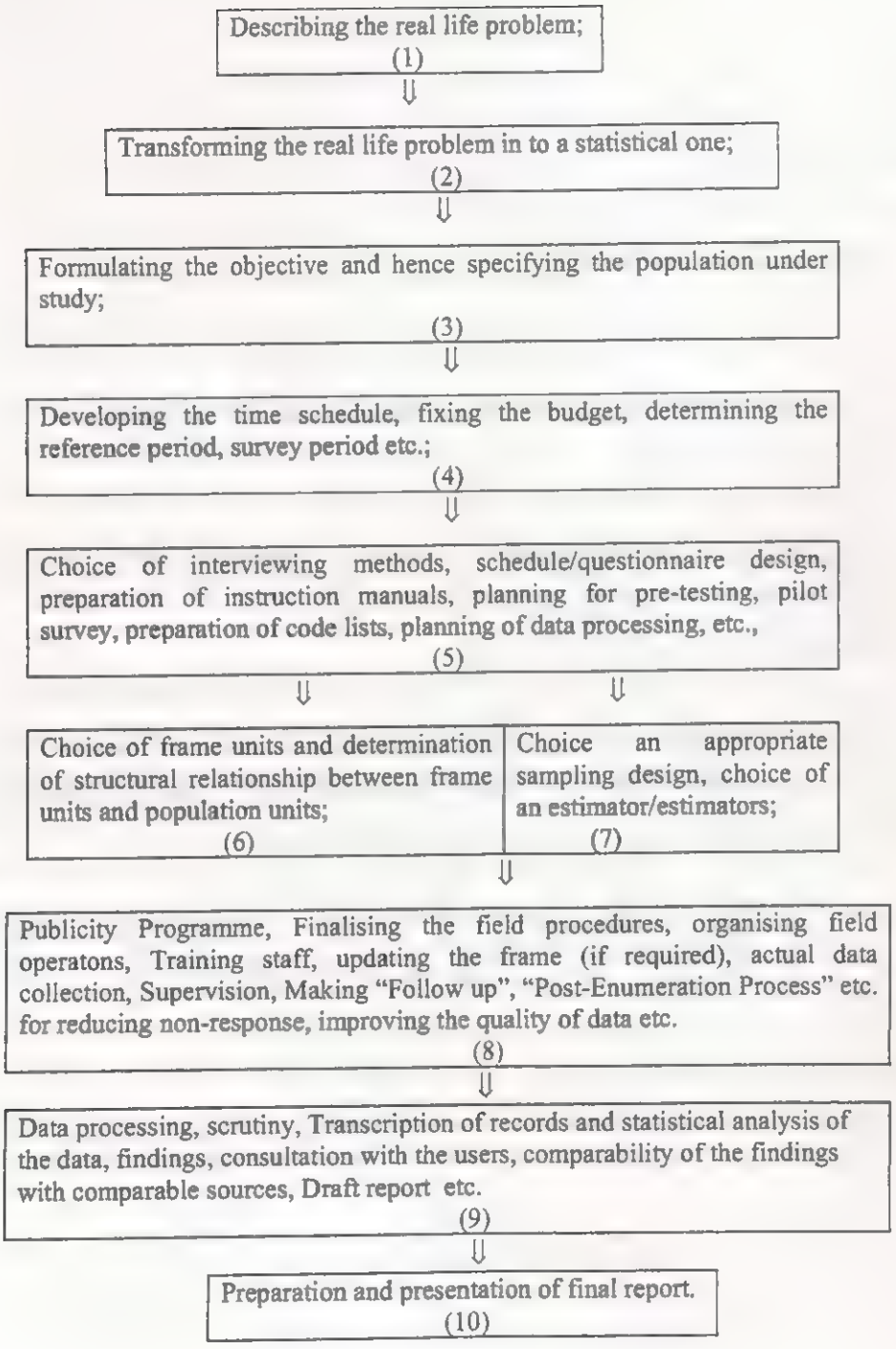
A survey may mean to include 'census' which attempts to study all members of population, where as a sample survey refers to a survey in which a scientific sample of the populations is studied. One may argue that non-probability samples are in some instances are scientific [REC (1975), survey of family and community life in South Africa, (1995), IPP-VIII Project (1998)]. Sometimes combination of probability and non-probability samples may be used. But in this presentation, we choose to exclude from our consideration studies using non-probability samples.

2. Major steps in the total Survey Design:

One can identify different activities associated with the sample survey procedure in the total survey design through the following major steps. It is true that each survey is unique in its own features, but even then the following schematic diagram may help some one work within a given set up and help in making assessment of

(a) total time needed with its break-up at various stages;
(b) cost components at various stages;
(c) an appropriate estimator, and
(d) the requirement of the format of tabulation programme, etc.

This will also help in understanding the different pockets in the total survey design which may act as different sources of errors both -- sampling and non-sampling, which are likely to occur in the integrated system of the survey operation.


Describing the real life problem;
(1)
Transforming the real life problem in to a statistical one;
(2)
Formulating the objective and hence specifying the population under study;
(3)
Developing the time schedule, fixing the budget, determining the reference period, survey period etc.;
(4)
Choice of interviewing methods, schedule/questionnaire design, preparation of instruction manuals, planning for pre-testing, pilot survey, preparation of code lists, planning of data processing, etc.,
(5)
Choice of frame units and determination of structural relationship between frame units and population units;
(6)
Choice an appropriate sampling design, choice of an estimator/estimators;
(7)
Publicity Programme, Finalising the field procedures, organising field operatons, Training staff, updating the frame (if required), actual data collection, Supervision, Making "Follow up", "Post-Enumeration Process" etc. for reducing non-response, improving the quality of data etc.
(8)
Data processing, scrutiny, Transcription of records and statistical analysis of the data, findings, consultation with the users, comparability of the findings with comparable sources, Draft report etc.
(9)
Preparation and presentation of final report.
(10)

The above activities need not always follow the same sequence as mentioned, but they should be considered in the whole of survey operation, at one stage or the other.

Non-sampling errors may be defined as a residual category i.e., all errors of estimation that are not the result of sampling. Thus one can have non-sampling errors arising from

(a) deficiencies in the problem formulation;
(b) inability to include a frame unit and linking up with a population unit;
(c) improper designing of schedules/questionnaires;
(d) faulty method of data collection;
(e) lack of proper training of field and supervisory staff;
(f) inaccuracy in data processing etc,.

The earlier work of assessing and classifying non-sampling errors started during the fourties [Deming (1944) (1950)]. At the 32nd session of the International Statistical Institute, a paper on non-sampling errors was presented by Mahalanobis and Lahiri which appeared in Sankhyā (1964). Among others, the works of Mahalanobis[(1940),(1944),(1946)], Moser (1958), Zarkovich (1966), Dalenius [(1977a), (1977b), (1977c)], Cochran (1977), Kish (1965) and Samdal et al (1992), may be mentioned in this direction.

We concentrate here, among all other different types of non-sampling errors, only the mechanism which produces a response or by which non-responses occur. Various sources of non-response are discussed. The relative importance of these factors and the measures needed to control them will vary from country to country, from one culture to another and from one survey to another, as each survey is characterized by its own distinctiveness with respect to the objective, the population understudy etc. In any situation, the objective of identifying the major sources of non-response, and the characteristics of the field staff and respondents associated with non-response is to devise measures to control non-response, to adjust for it and to estimate its effects on survey results.

Organization of this presentation is as follows:

To begin with, a formal definition of a response (unit and item) / non-response is presented, followed by some illustrations on the existence of non-response from surveys in both developed and developing countries. They exhibit characteristics of who would be set of non-respondents and who not. As we shall see from the illustrations that in general, refusals to participate in the survey are an important source of non-response in developed countries, but a relatively minor component of non-response in developing country household surveys. Finally, we discuss some methods dealing with non-response.

2.1. Probability Sampling and Non-Response:

A practical limitation to the probability sampling in survey research is that the sample actually achieved suffers defficiencies from non-response and non-coverage. Both of them can have bias in the survey estimates. Non-coverage occurs when certain elements in the target population are not included in the sampling frame from which the sample is selected, whereas non-response as a concept has been defined in a number of ways. Most definitions distinguish unit non-response from item non-response. In general non-response has been attributed to failure to obtain a response to a particular unit and/or to a particular item when the questionnaire has been completed. [Kendall and Buckland (1960), Kish (1965), Bureau of Census (1957, 1976), Cochran (1977), Zarkorich (1966), ford (1976), National Research Council (1983), Sudman (1976), Suchman (1962), Wark and lininger (1975), Deghton etal. (1978), Deming (1953) etc.,]. Thus, non-response depends at all the stages of the integrated system of total survey design, -- contrary to the general belief that it occurs only at the interactive process between a respondent and an investigator. We shall provide in the next section the causes of a unit non-response as well as an item non-response.

An extended definition of non response includes in which missing data arise

(a) from the processing of information provided by units rather than refusal of units to provide information. For example, editing procedures may eliminate some responses which are to be judged to be impossible and inconsistent with other findings.

(b) out of the problem of non-contact due to inaccurate assessing information to reach a sample unit, - due to inappropriate responding rule, inappropriate linking rule between the target element and the frame/survey unit.

(c) because of non-availability (temporarily) of a respondent at home for the inappropriate choice of time of interview.

(d) because of non-coverage due to inaccurate frame.

(e) for lack of solicitation to make the respondents participate in the survey process (refusal).

(f) due to difficulties in contacting for natural calamities like floods, earthquakes and / or for political disturbances.

3. A Review of a catalogue of problems encountered at different stages in some Surveys conducted both in developed and developing Countries.

The refinement of the general data requirement of any sample survey into the precise questions is a step-by-step process and needs resolving certain issues in the conceptual frame work for the translation of the real life problems into a statistical one. Failing to this, the target population cannot be unabmbiguously defined, and hence coverage errors are likely to occur. Any kind of non sampling errors will have their influences on the response and/or non response.

A response or a non response is the outcome of an interactive process between a respondent and an investigator working under a given survey condition. A measurement or response error occurs when incorrect value/judgement is associated with a population element and must not be taken to mean that such errors are the fault of the respondent alone. In fact, the outcomes of the process of data collection, namely, unit/total non response, partial response, multiplicity of the same response, response with or without error depend on all persons who take part in the total survey operation. The failure to obtain qualitative data depends on many factors such as,

(a) the inaccuracy in the definition of the population units [Domestic Tourists Survey in Orissa (1988-89)];

(b) making frequent changes and having non uniformity in the definition [Report of National Statistics Commission, August, 2001].

(c) non-availability of a frame [Domestic Tourists Survey, Orissa, India (1988-89); Stanza Bopaper Project, Mamelodi, South Africa (1995-96); Calcutta Urban Poverty Survey (1976)].

(d) the inaccuracy in the frame;

(e) the inaccuracy of survey materials (length of questionnaire, wording and ordering of the question, length of recall period, instruction mannul etc.);

(f) inadequacy in uniform training of field staff and negligence on their part;

(g) lack of proper training of interviewing techniques [Platek (1977), Fellegi (1963), Thomsen and Singh, Cole (1965)]

(h) interviewer deficiencies (poor interviewing techniques, misunderstanding of concepts, misinterpretation of response, wrong arithmetic etc.,; his gender, employment status, ability to create rapport etc.);

(i) difficulty in implementing a random sample due to a peculiar field condition,

(j) respondent's failures due to interpretation of the question, inability to provide answers and deliberate or inadvertent supply of wrong information, etc. and also their preferences for some numbers [1976 fertility survey in Indonesia, Dasgupta and Mitra (1958)];

(k) inappropriate respondent rules [Tuigan and Cavdar (1975)]

(l) imposition of social stigma like those of female participation in work force [Shah (1981)], taking alcohols etc.;

(m) purposeful reporting of certain information incorrectly, such as women do not like to disclose their ages etc.;

Unit non response may occur for any or some of the following reasons:

(a) The problem of non-contact due to inaccurate assessing information to reach the sampling units;
(b) The problem of non-availability at home (temporarily) because of inappropriate choice of time of interview;
(c) The problem of non-co-operation i.e., the problem of refusal;
(d) The problem of communication between the data collector and the sample member;
(e) The difficulties arising out of natural calamities like floods, earthquakes etc. and of political disturbances;
(f) The problem of non-coverage;

Causes of item non-response:

- Identification of the items which produce higher non-response rates, specifically sensitive items such as income etc., [Donald 1960];
- Mode of interview is responsible for producing item non-response;
- Participant's characteristics affecting item non-response [Ferver, (1966)] higher items non-response rates arise on questions enquiring substantial thought or effort on the part of the respondent. [Frances and Busch (1975), Craig and MC-Cann (1978)];
- A significant age [Messmer and Seymour (1982)] and occupation effect, while noting that the extent of item non-response did not seem to depend on questionnaire length [(Craig and MC-Cann (1978)];
- Messmer and Seymour (1982) discovered that questions appearing after a branching question had notably higher item non-response rates than did other questions;
- Rogers (1976) found interviews who were more impersonal to have lower item non-response rates than those of interviewers who had a more personable interviewing style;

- Sudman etal. (1977) discovered that interviewers who entered the interview thinking that the questionnaire would be difficult to administer had higher item non-response rate than did more optimistic interviewers;
- Bailer etal. (1977) reported that interviewers who thought it inappropriate to ask a sensitive question had higher item non responses on the question;
- Non response on some items will be higher for some subgroups (elderly, females, the less educated);

4. Effects of the foregoing factors on the out come of the data gathering operation-response/non response:

It has been empirically observed through a number of surveys [(Bennet and Hill (1964), Cobb, King and Chen (1957), Dunn and Hawks (1966), Lubin, Levitt and Zuckerman (1962), Lundberg and Larsen (1949), Newman (1962), Ognibene (1970), Pan (1951), Reuss (1943), Skelton (1963), Warwiek and Lininger (1975), Kendal and Buckland (1960), Madow et all (1983), U. S. Bureau of Census (1974), Kalton (1983), Sudman (1976), Suchman (1962), Sukhatme (1970), Birbaum and Sirken (1950), Deighton et all (1978), Politz and Siman (1949), Madow et all (1983 (a)), Gower (1979), Demaio (1980) Kalsbeek and Lessier (1978), Lessier [(1974), (1980)] Roy [(1976-77), (1977-78),(1988-89), Maiti (1994-95), (1995-96)] etc.
that

(a) non-respondents differ with respect to income class; People with higher incomes refuse more frequently, though no difference has been observed with respect to race or sex [(Demaio 1980)] ;
(b) non-respondents differ with rural and urban composition; [(REC-Project (1975-76)];
(c) non-responding households differ from responding units in household size and labour force status [(Gower, 1979,)];
(d) non-respondents differ from respondent by tennure status. Owner of houses has a refusal rate 14.1%, where as for local Council tenants, the rate was 5.3%.(Barnes and Birch, office of population census and surveys, Study NM1, London);
(e) non-respondents cannot be characterised by sex or race; although female respondents co-operate more to the female investigators;[(Domestic Tourists Survey (1988-89), IPP –VIII (1998)], especially incase of sensitive questions;
(f) older people and people with higher income constitute a set of non-respondents (Demaio, 1980);
(g) non-response rates vary by different reasons, socio-economic groups etc.[UK general household survey (1971-74), Lyberg and Repaport (1979), UN commission on employment and unemployment statistics (1979), Norweign Labour Force Survey (1972-78)]
(h) non-respondents vary with respect to different choices of 'recall period';
(i) effect of non-response will be higher if the set is very much dissimilar in respect to the variable under study;
(j) non-response rate is gradually increasing over the year[unpublished work of Thomosen and Siring]
(k) non-response rate due to refusal is greater than due to non-contact in

developed Countries, where as in developing Countries the reverse picture holds [Verma (1980), Norweign Labour Force Survey (1972-78)];

(l) non-response rate varies from 15% to 46% [United Kingdom, office of population censuses and surveys (1978) IPP-VIII(1998)];

(m) according to the study conducted by the Sweedish Central Board of Statistics (Bergman et al, 1978), refusal takes place because (i) people are very much concerned about privacy, confidentiality, (ii) some respondents fail to understand a random sample and decline to give information.

(n) non-responses occur due to coverage errors [Hirschberg etal.(1977)] and coverage errors may be classified according to one of following different types.

(i) Coverage errors due to inclusion of non-population and /or exclusion of population units; In a recent study over coverage was found to be around 15% [IPP –VIII population project (1998)];

(ii) Coverage error due to smaller size and missed structure. An analytical study of the results of a report by Turner et al (1979) on "1976 national survey of farm production in Domain Republic", led to the conclusion that the survey estimate of the number of farms was low by 15 to 22%. A useful study of non-coverage of dwellings was provided by Kish and Hess (1958). Similarly, an analysis of the data from US monthly labour force survey (the current population survey) by Brooks and Bailar (1978) reveal that less than 3% of the target population was not included because of missed structures;

(iii) Coverage error due to faulty selection procedure may arise from incorrect application of the sampling scheme [1951 census of live stock in Yugoslovia; Zarkovich (1966),].;

(iv) Coverage error due to element multiplicity: As desired by the National Advisory Board of Statistics (NABS), a sample survey scheme was drawn as a possible replacement for a national Economic Census (E.C.) [Ghosh et al (1999)]. It was observed that in the village a single person might be engaged in a number of small enterprises. While using the list of such enterprises as frame units in estimating the total number of workers, multiplicity of an worker might have lead to over estimation of number of workers, if an appropriate method of estimation was not adopted;

(v) Coverage error due to failure of identifying and contacting a population unit because of inaccurate assessing information associated with the frames (IPP-VIII Population Project (1998)], Tuygan and Carvador (1975)] ;

(vi) Extreme case of coverage errors was reported for the 1946 census of Industrial and Business establishments in France (Chebry, 1999); other examples of coverage error can be found in the works of Chapman and Rogers (1978), US Bureau of Census (1978), Sample census of population in Wales and England (1966), Gray and Gee (1972), Demographic Year book of the U.S. (1956) census of Srilanka (1953), Population Census of Liberia (1980);

(vii) Errors due to outdatedness of the frame may be those which occur when information about the element does not permit location; [Platek (1977), Fellegi (1973)]. Recently such a situation did occur when the ISI alumni were required to be contacted;

(viii) Duplication of the frames: causes coverage errors. One can observe duplication of units in the frame lists maintained by the office of the DCSSI and CSO/NSO in conducting Annual Survey of Industries (ASI) [(Report of the National Statistics Commission, August 2001), Kish (1965)];

(o) inappropriate design/format of the questionnaires/schedules cause problems. As an example, one can find that the questionnaires/schedules available with the office of the DCSSI, in conducting Annual Survey of Industries (ASI), is not in consistent with type of data required [Report of the National Statistics Commission, August, 2001)]

(p) inordinate length of the schedule/questionnaires increases the respondent burden causing recall lapse error. To avoid such error, use of a matched sample is sometimes made [REC-project (1975-76), IPP-VIII (1998)]. Multiple reference period are also used to reduce respondent burden [(NSSO: 51st through 54th and 55th rounds of survey)]. Respondent burden can also be reduced by administering different questionnaires to different respondents in a common sample of areas (Rao and Shastry, 1975), to different rounds of the same survey (Domestic Tourist Survey, (1988-89).

(q) choice of appropriate length of the reference period should be made by considering "heaping effects", "blocking effects" etc. and also the sampling error of the estimate should be taken into consideration while preparing the reference period. The period should be neither too small nor too large. Displacement error is another type of error due to inappropriate length of the references period and can be avoided by making the reference period closed or bounded. [Mahalanobis and Sen (1954), Scott (1973), Ghosh (1953), Neter and Warsberg (1965)];

(r) It is not unlikely for some of the investigators will tending to favour small households to keep the work load small; others with good intention may substitute larger neghbouring households for small households scattered. That filed conditions pose a serious problem in executing a random sample could be seen from the field work in connection with the data collection of the REC-project (1975-76). All the electrical poles were broken due to heavy storm on the day when we reached a village in the district of Burdwan for data collection, The villagers thought us to be State Electricity Board (SEB) personnel and every body came forward for providing information. Selecting a few of them through a probability sample and interviewing a part became almost impossible;

(s) there may be problem on non availability of imputed values: one can experiences, while compiling indirect estimates of Gross Value Added (GVA) in the formation of national account statistics in India, that imputation of many supporting values may not be readily available [Report of National Statistics Commission, August, 2001];

(t) one can commit mistakes in transcription and recording information at the data processing stage [REC project (1975-76), Domestic Tourists Survey (1988-89), IPP-VIII (1998)];
(u) response bias arises out of Panet data also [Bailar (1979)].

5. Classification of the previous non-sampling errors (Non-response errors):

Classification procedures adopted may be of any of the four types. They may be classified according to (i) where they occur or (ii) according to some specific characteristics of the population under study or (iii) according to the type of measures--preventive and/or compensatory by which they can be controlled, or (iv) according to the nature of inbuilt stochastic elements.

A. Classification according to the places of occurrence:

A1. During planning or Pre-data gathering stage:

(i) Errors of coverage due to inclusion of non-population and exclusion of population units for using a faulty frame [n(i), n(ii), n(vi), of Section 4];
(ii) Errors due to elements' multiplity [n (iv) of Section 4];
(iii) Errors in the frame due to faulty cluster size, affecting the inclusion probability [n (i) of Section 4];
(iv) Errors due to non-contact for incorrect assessing information; [(k), n(v), of Section 4];
(v) Errors of recall lapse, of displacement due to ill designing of the schedules/questionnaires [(h), n(o), (p), (q) of Section 4];
(vi) Errors of coverage, due to choice of an inappropriate sampling design as well as a sampling frame [n(iii), (r), (u), n (vii), n(viii) of Section 4];
(vii) Errors due to faulty method of selecting the investigators [(e), of Section 4];

A2. During data gathering stage.

(i) Total and/or partial non-response;
(ii) Response or measurement errors;
(iii) Rotating panel bias.

A3. During post –data gathering stage.

(i) Processing errors [(t) of Section 4];

(ii) Errors due to rounding off the weights/multipliers meant for inflating the estimator for the population parameter [(s) of Section 4];

(iii) Tabulation errors.

B. Some specific characteristics of the population under study sometimes appear to be responsible for causing response and/or non-response errors [(a), (b), (c), (d), (f), (g), (m), of Section 4].

C. Errors of Preventive as also of compensatory measures:

The above different types of errors can also be alternatively classified in following errors of preventive and / or compensatory major.

The preventive measures are those that would be implemented for identification, solicitation and compilation of the questionnaires, so that after the sample member has agreed, at least, in principle to co-operate, relevant data can be made available smoothly.

Despite one's best effort for taking preventive measures to minimise errors, there will be still possibility of facing problems during and /or after data collection, measures of which can be termed as non-preventive or compensatory measures.

A particular type of error can be identified as belonging to the class of errors of preventive as also those of compensatory depending on the type of measures one has taken against that particular kind of error.

C 1. An illustrative list is shown below.

(1) Frame errors can be defined as errors of preventive ones, if some or all of the following actions are taken.

(a) Adoption of an appropriate frame of all materials which will describe the components of the target population (or an adequate part of that population) in

such a way that it will be possible to identify and interview individual components during the course of the survey operation.[Zarkovich (1966), Scheaffer (1970), Hansen et al (1963), Szameitat and Schaffer (1968), Jessen (1978) etc].

(b) Defining an appropriate association rule between the frame unit and the population unit;

(c) Identifying the appropriate observable unit. In a multipurpose household survey, there could be more than one responding units within the same household [IPP – VIII (1998)];

On the other hand, if incomplete or erroneous frames are used for collecting data and then appropriate measures are taken at the estimation stage, then such errors due to the use of an incomplete or an erroneous frame can be termed as errors of compensatory measures;

(2) Investigator variation can be termed as an error of preventive measure, if the following action is taken.

(a) providing uniform training of the interviewing techniques. [Graves and Kahin (1979), Ferber (1966), Frances and Bush (1975), Craig and MC Caun (1978), Messmer and Scymor (1982), Ford (1968), Sudman (1977)];

(3) Response error can be of error of preventive measure if the following technique is used

(a) Use of randomised response technique (Warner, 1965) to reduce non response rate; Several adaptation of Warner's original idea are available [Horvitz et al (1967), Greenberg et al (1969), Folsom et al (1973), Emrien (1983).

(4) Non response errors-total and/or partial can be of preventive type,

if some or all of the following acts are made.

(a) Choice of appropriate structure of the questions and of appropriate reference period at the designing stage is made so that respondent burden and hence errors

of "recall lapse" and "displacement" due to the respondent burden can be reduced;

(b) Planning is made for the use of a matched sample;

(c) Provision is there for multiple reference periods;

On the other hand, non-response errors can be thought of errors of compensatory measures, if some of the following methods are adopted during and/or after the period of data collection.

(a) Method of call backs: [Cochran (1977), Deming (1953), Kish (1965), Politz and Siman (1949), Birbaum and Sirkan (1950), Deighton et al (1978), U.S. Bureau of Census (1975), Kendal and Buckland (1972), Moser and Kalton (1972), Durbin (1954); Roy (1975-76), (1988-89)];

(b) Proxy Interview: [Roshwalb (1982), Roy (1975-76)];

(c) Intensive follow up: This method is proposed specially in connection with mail surveys; Intensive follow up may be based on the deterministic view of formation of strata of the respondents and the non-respondents [Hansen, Hurwitz (1946). Fellegi and Sunter (1974)], or may be based on the stochastic view of the strata formation of respondents and non respondents [Platek et al (1977)];

(d) Substitution: New sample members in this approach are substituted for unit non respondents as a means to maintain the intended sample size although the bias from non-response will not be reduced [Roy (1975-76), Maiti (1994), (1998)];

(e) Prediction Approach: Information on non respondents can be obtained through information on the respondents following both Bayesian and non Bayesian model based approaches. Both the methods assume a distributional assumption on the study vector, say, $\underset{\sim}{Y}$, on the selection process D which is binary and also on the response vector R which is also binary [Ghosh (1997)]. The use of model based methods to deal with survey response in particular and for survey inference in general have been summarised by Hansen et al (1983). Works due to Rubin (1977), Little (1983), Platek and Grey (1983), Shah (1981), Rubin (1978, 1979), Herzog (1980), Cox and Folsom (1981) can be referred in

connection with the model based methods for drawing inferences for the population under study.

(f) Imputation Techniques: Cross classificatory based imputation techniques includes cold deck procedure and hot deck procedure [Champan (1976), Nord botter (1963), Oh and Scheuren (1980), Cox (1980), Chromy (1979), Rizvi (1983), Cox and Cohen (1985)], where as model based imputation assumes a statistical model about non-response and about the formation of the sampled population. [Kalton and Kasprzyk (1982)]. Parameters of the model are estimated from the respondent data and the fitted model is then used to predict item value for the non-respondent.

(g) Adjustment for the weights: Adoption of one of the methods to adjust for non-response at the estimation stage includes that of any of the works due to [Politz and Simmons (1949)], Platek et al (1977), Kish and Anderson (1978), Hartigan (1975), Hansen et al (1953), Kohen and Kalsbeek (1981), Bailar et al (1978), Rizvi (1983), Madow (1983), Drew and Fuller (1980, 1981), Thomsen and Siring (1983), Maiti (1998)] etc.

5. Rotating Panel bias arises due to (a) acquaintances of the respondents with the structure of the questionnaire, (b) acquaintance with the interviewers, (c) some conditioning effects and (d) change of the characteristics of the respondents which the interviewer may apprehend and this panel bias depends on the number of times a panel is exposed to. [Bailar (1975) (1979)], Palan (1978), Mooney (1962), Woltman and Bushery (1975), Hansen et al (1955), William and Mallors (1976)]. This bias may be of preventive type if an appropriate sampling design at the planning stage is adopted. The sampling design adopted should be such that at regular intervals of time, a fixed panel may be dropped and replaced by new panel of the same size.

6. Dealing with non-response:

Several methods can be used to try to compensate for the effect of non-response on the results. Those can be associated with the different kinds of activities in the whole of survey operation. Broad classifications of the methodologies involved depend on whether they are at some or all of the following stages:

(A) Pre-data collection / planning stage;

(B) During field work;

(C) Post-data collection stage.

(A) Preventive measures at Pre-data collection / planning stage:

(i) Specification of an appropriate frame unit and of a linking / counting rule between a target element and a frame/ survey element, is a must one without which there would not be proper specification of the population under study, causing coverage errors – under or over.

(ii) Non-response among the data items obtained from n_1 respondents from the original sample can be controlled by a strategy which calls for preventive measures. These serve to reduce the chance that participants will fail to respond or provide useful data for individual questionnaire item. For this, one must understand what qualities of the survey contribute to increasing the incidence of missing data for some questionnaire items. Here arises the question of designing good schedules/questionnaires.

Apart from the basic need of a sampling design to provide a sound basis for drawing statistical inferences, the other important factor to consider in any survey is the questionnaire design. If the sampling design is intended to provide a representative sample, questionnaire design is an important means of collecting data on the selected units. The quality of the data is as important as the sampling design or even more. After determining the concepts, definitions, classifications to be used, the next important step is the preparation of a detailed list and description of the survey variables with their units of measurement, before they are intended to be presented in a most efficient way as a data gathering instrument. The variables to be measured have to be transformed into operational definitions expressed in the form of a logical series of questions which the interviewer can ask and the interviewer can comprehend and answer. They should be designed in such a way that they do not suffer any kind of defficiencies and

- enable the collection of accurate information;
- facilitate the work of data collection, data processing and tabulation;
- ensure economy in data collection;
- permit comprehensive and meaningful analysis and purposeful utilization of captured data.

These inturn needs to decide

- whether the schedule should be a structured or a non-structured one;

- order of placement of the items of information; consideration should be given regarding sensitive and / or non-sensitive questions;
- on the use of local dialectics;
- on mode of data collection; Interview surveys conducted by telephone tend to have more missing items than those conducted in person (Groves etal. 1974);
- on the length of the reference period [Mahalanobis and Sen, (1954); Scott (1973), Ghosh (1953)];
- on how to reduce recall lapse/memory strain [Donald (1960), Maiti etal. (1998)];
- on who would be the respondents [Rogers (1976), Bailar etal (1977), Sudman etal. (1971), Ferber (1966), Francis and Bush (1975))];
- on the questionnaire length [(Craig and MC Cann (1978), Messmer and Seymour (1982), Ford (1968)], etc..

(iii) Interviewer Aids and Accompanying Documentation:

In the implementation of the survey, the questionnaire/ schedule needs to be supported by various types of other documents and aids so as to facilitate the collection of data without much error as far as possible. They may be as follows:

A manual of training: this is primarily aimed at the trainers and supervisory staff including those engaged in filed pre testing and evaluation of the questionnaire. The manual should cover topics such as

- objectives;
- organization and procedures of the training programme;
- identification of the role and training needs of every staff member at all levels;
- general content and timing of training for various operations such as the preparation and selection of the sample, pre testing, recruitment and selection of field and other staff, field supervision, interviewing, editing and scrutiny, coding and data entry.

An instruction Manual for Interviewers:

The interviewer instruction manual is the most critical document accompanying the questionnaire, since interviewers are the personnel most directly concerned with

implementation. The manual should help them understand their duties and provide instructions on procedures and techniques. It should also provide a detailed explanation for each of the questions and clarification of concepts and definitions.

Interviewers should be taught of the following techniques.

- The interviewer should win the confidence of the respondent;
- Self-introduction to the respondent (ID card letter etc);
- They should make the respondent understand what is needed and the purpose of the survey;
- They should use simple language, preferably the local dialect;
- Questions should be asked exactly as worded in the questionnaire;
- The interviewer should not interrupt the respondent;
- No answer should be assumed;
- There should be no overlapping between asking any two questions;
- Only the selected persons should be interviewed as far as possible;
- Interviews must not be conducted when outsiders are present.

B. Methods Adopted during Field Work:

These methods are part of the data collection procedure, for example, intensive follow up of a sub-sample of non-respondents, or the collection of limited data from neighbors for households that are away during the data collection period. The substitution of other units for those units which can not be interviewed is a controversial practice. However, non-response in household surveys can not be effectively dealt with unless it is properly identified during data collection. It should be standard practice for interviewers to account for the outcome of every sample unit assigned to them. This means recording whether or not an interview was obtained, and if not, they should explain the circumstances in sufficient detail so that each unit not interviewed can be classified as,

- eligible for interview;
- not eligible;

or • eligibility not determined.

This information should be transmitted to the data processing unit for use.

B1. Method of call backs:

The recommended way of dealing with non-response during the data collection stage of the survey is to make a vigorous and thorough effort to obtain responses for all,

or nearly all of the eligible units in the assigned sample. If no acceptable respondent is available, when a unit is first visited, call backs should be planned. If possible, the neighbour should be asked when the occupants are likely to be at home. The optimum number of call-backs in a particular survey depends on several factors. The following tables provide us empirical evidences for effect of extra call-backs.

Table 6.1: Effect of Extra call backs on non-response rates in Norweign Labour force survey: 1972 – 1978.

Year (data are for second quarter)	% non-response rate		% distribution by reason		
	Normal procedure	After extra call backs	Refusals	Not at home	Others
(1)	(2)	(3)	(4)	(5)	(6)
1972	8.4	7.3	53	39	8
1973	10.5	9.5	NA	NA	NA
1974	10.6	8.8	41	42	17
1975	11.1	7.4	45	43	12
1976	14.5	9.9	45	38	17
1977	14.1	10.4	47	37	16
1978	16.4	11.9	45	42	13

Table 6.2: Interviews completed, by number of calls for a household expenditure Survey in Great Britein*.

Number of call	1	2	3	4	5	6 & more	Total
Additional % of households connected	62.0	22.3	9.0	4.1	1.6	1.0	100.0

* The data in Table 6.2 is due to Cole (1956)

Deming (1953) developed a useful and flexible mathematical model for examining in more detail the consequences of different call-back policies. He also examined the cost effectiveness of call-backs and showed that in some situations a large number of successive calls may be justified.

According to him the population is divided into *r*-classes depending on the extent of the probability with which the respondent will be found at home.

The true population mean $\mu = \sum_{j=1}^{r} p_j \mu_j$, where p_j = proportion of the population falling in j^{th} class.

Let w_{ij} be the probability that a respondent from the j^{th} class will be reached on or before the i^{th} call. Let n_0 be the initial sample size and $\bar{y}_i$ be the sample mean obtained after *i* calls.

Then, $E(\bar{y}_i | n_i) = \sum_j \left(\frac{w_{ij}\, p_j}{\sum_j w_{ij}\, p_j} \right) \mu_j$ *i.e.*, $E = (\bar{y}_i) = E(\bar{y}_i | n_i) = \sum_j \left(\frac{w_{ij}\, p_j}{\Sigma w_{ij}\, p_j} \right) \mu_j$

For dealing with "the not-at-home" cases, a procedure that saves the cost of call-backs which consists of ascertaining from the respondents the chance of their being at home at a particular point of time may be adopted. Suppose, the enumerators make calls on households during the evening on 6 nights of the week. The households were asked whether they were at home at the time of interview on each of the preceding nights. The households may then be classified according to *r*, the number of evenings, they were at home out of five and the ratio $(r+1)/6$ is taken as an estimate of the probability of a household being at home (Poletz – Simmon 1949). Let n_r be the number of interviews obtained in the group *r* and $\bar{y}_r$ is the group mean, then the estimate of the universe mean would be

$$\hat{\mu} = \sum_{r=0}^{5} \frac{(n_r\, \bar{y}_r)\, 6/(r+1)}{\sum_{r=0}^{5} n_r\, 6/(n+1)}$$

B 2. Proxy Respondents:

A proxy respondent may be defined as person, who although not selected in a sample is considered a suitable substitute, when the person actually selected can not participate. When repeated call-backs fail, it may be possible to partially complete a questionnaire for the assigned unit by observation (e.g. for many housing characteristics) and by asking neighbours for information. Interviewers should be instructed which of the

survey items they may ask of neighbours. These items should include only non-sensitive questions for which neighbours might be expected to give reasonably accurate answers. By using proxies, non-response will be less, but they may contribute to increased measurement error, since facts or opinions may be less accurately portrayed [Roshwalb (1982)] .

6.3: Effects of Proxy Response on Demographic Items: Pilot test for Turkish Demographic Survey.

Item	Urban area		Rural area	
	Procedure 1*	Procedure 2*	Procedure 1	Procedure 2
% of women responding themselves	88	75	66	59
Average number of children ever born to ever married women	3.00	2.81	5.36	5.07
Proportion of children dead at time of Interview	0.207	0.178	0.266	0.254

* Under Procedure 1: "Interviewers were asked to make special efforts to obtain responses from all eligible women to questions relating to their own children".

* Under Procedure 2: "No such as above instructions were given and proxy respondents were used more often".

B 3. Non-respondent Sub-sampling

For a number of years, survey statisticians favoured a deterministic view according to which the population is dichotomized into a response and a non-response stratum (see section 4). By assuming that members of the population are either certain to respond $(p_i = 1)$ or not respond $(p_i = 0)$, the deterministic view of non-response removes any uncertainty as to whether or not, each member of the population would provide useable data for the survey, if selected.

Following a reasonable effort to obtain response by standard survey procedure, a sub-sample of the remaining non-respondents should be selected and more intensive

procedures are used to obtain the survey information for these units. This method is generally attributed to Hansen and Hurwitz (1946). However, Dalenius (1957) points out that the idea was first suggested by Cornfield (1942).

(1) A sample s of size n is drawn, in the first phase, according to the design p_s with positive inclusion probabilities Π_k and $\Pi_{k\lambda}$ for k^{th} unit and for a pair (k,λ).

(2) Despite efforts to obtain responses y_K for all $k \in s$, some non-response occurs. The sample s is dichotomized into a response set s_1 of size, n_1 and a non-response set s_0 of size n_0.

$$s = s_1 U s_0; \quad n = n_1 + n_0$$

(3) A suitable large sub-sample s_0' of s_0 with $n_0' (< n_0)$ number of non-respondents is drawn by a design $P(\cdot \mid s_0)$ with positive inclusion probabilities $\Pi_{k|s_0}, \Pi_{k\lambda|s_0}$. The necessary efforts are made to record a response from every element.

The requirement of full response may prove costly, but it is the requirement that makes unbiased estimation possible.

An unbiased estimator of the population total is given by

$$\hat{y}_{H-T} = \sum_{K \in S} y_K / \Pi_K^*; \; s = s_1 U s_0'$$

$$\text{where, } \Pi_K^* = \begin{cases} \Pi_k \text{ if } k \in s_1 \\ \Pi_k \Pi_{|s_0} \text{ if } k \in s_0 \end{cases}$$

with appropriate variance (i.e., total variance),

$$V(\hat{y}_{H-T}) = E_1 V_2 (\hat{y}_{H-T}) + V_1 E_2 (\hat{y}_{H-T}).$$

> Much of the recent work on estimation in the presence of non-response is based on the idea that the response is stochastic, but not deterministic. Under such an outlook, a response distribution is assumed to exist. One might reasonably argue that if the same protocall for location, solicitation, and data collection were repeated, a sample unit might respond on one occasion and then not respond on another occasion, [Platek etal. (1977)].

Introducing a stochastic variable on the outcome, the above estimator can be revised *i.e.*, there exists a response mechanism / response distribution (RD) that governs the dichotomization of the samples s into one responding subset s_1 of size n_1 and non-

responding subset s_0. This implies that if a given s were surveyed repeatedly, the composition of subsets would vary from one survey to the other.

Let $P(k \in s_1) = p_K$ and $P(k \in s_0) = (1 - p_k)$.

Using this information, the estimator

$$\hat{y}_{H-T} = \sum_{K \in S} y_K / \Pi_K^*$$

can be revised.

B 4. Substitution: substitution of other units for the non-responding units has been used in several household surveys, both in developed and developing countries. The rationale for its use is usually to ensure that completed interviews will be obtained for the exact number of sample households specified in the initial design.

Substitution does not eliminate non-response bias. This may be understood if one views the survey universe as being divided into two groups: those households for which it is possible, by following specified survey procedure, to obtain an interview, and those for which, using the same procedure, it is not possible to obtain an interview. Substitution increases the sample size for the first group, but does not provide any representation of the households in the second group. Characteristics of the two groups are certain to differ (see section 4) and the substitution process has done nothing to reduce bias resulting from these differences. Substitution does control sampling error by achieving the desired sample size.

The main argument against substitution is that

- Frequently the rules established for substitution are biased;
- It is extremely difficult to prevent interviewers from making unauthorized substitutions;
- The use of substitution diverts attention from the problem of non-response bias.

B 5. Improved Mechanism of Data Collection:

A technique introduced by Warner (1965) assumes "the probability P with which the choice device selects the statement S" to be known to the statistician.

$$\text{Let} \quad y_K = \begin{cases} 1, \text{if } k^{th} \text{ individual in the finite population} \\ \quad \text{of size } N \text{ has the attribute } A, \forall k \in \mathrm{Y} \\ 0, \text{ otherwise.} \end{cases}$$

Let, $s = s(i_1, i_2, \mathrm{K}\ i_n)$ be a sample and for $k \in s$

$$x_K = \begin{cases} 1, \text{if } \textit{the } k^{th} \text{ individual gives the "true" answer,} \\ 0, \text{ otherwise.} \end{cases}$$

Let R – C indicate "with respect to randomized choice device".

$$E_{R\text{-}C}(x_K) = Py_K + (1-P)(1-y_K)$$
$$= (1-P) + (2p-1)\,y_K$$

thus, $\hat{y}_K = \dfrac{x_K + p - 1}{2p-1}$ and is unbiased for y_K.

and the variance is

$$V_{RC}(\hat{y}_K) = \frac{P(1-P)}{(2p-1)^2} = v_0, \text{ say;}$$

Let $\hat{y}_{PR} = \sum_S \hat{y}_K \Big/ \Pi_K$ be an estimator for population total, then

$$E(\hat{y}_{PR}) = E_p\, E_{R\text{-}C}\left[\sum_S \hat{y}_K \Big/ \Pi_K / S\right]$$
$$= E_p\left(\sum_S y_K \Big/ \Pi_K\right) = \sum_{K \in U} y_K = \text{Total.}$$

Using customary conditioning argument, we have,

$$V(\hat{y}_{PR}) = V_p\left[E_{R\text{-}C}(\hat{y}_{PR})/S\right] + E_p\left[V_{R\text{-}C}(\hat{y}_{PR} \mid S)\right]$$

Remarks 1: The second variance component $\left(\sum_u 1 \Big/ \Pi_K\right) V_0$ can be viewed as the price to pay for randomizing the response. The price may well be worth paying to get an unbiased estimate. The second component is large if P lies near $1/2$. So P should be chosen well always from $1/2$.

C. Methods Adopted During Data Processing Stage.

Both sub-sampling of non-respondents and randomized response require special arrangements that may be costly and time consuming. Especially in large scale standardized production of statistics, such arrangements may be impossible. Consequently, the survey statistician must often accept that some non-response is inevitable, and from the data that are actually collected, the best possible estimates must be produced in an efficient standard fashion.

Thus procedures used during data processing stage are generally less costly. These come under the general headings of "imputation" of missing data and "estimation procedure" which attempt to compensate for missing data. In general, the procedures used in data processing rely on assumption of similarity between responding and non-responding units, either in the whole population or preferably with more homogeneous subgroup of the population.

C1. Estimation Based Methods:

Typically, the weights assigned to sample data to produce estimates for the survey population have three basic components:

(i) Factors needed to adjust for non-response in the survey;

(ii) Factors reflecting the selection probabilities of the individual survey units;

(iii) Factors need to make estimated totals from the survey agree with comparable totals available from other sources.

Let us consider the simple problem of estimating a total Y, where we know the estimator due to Horvitz-Thompson (1952),

$$\hat{Y}_{H-T} = \sum_{i \in S} W_i y_i$$

when applied to a probability sample of size n will be unbiased. In the absence of (non-sampling error), $W_i = \Pi_i^{-1}$, where Π_i is the inclusion probability. When non-response is present and there are data for $n_1 (< n)$ sample members, the estimator $\hat{Y}_{H\ T}$ is not unbiased. In the presence of such error, an unbiased estimator for total would be

$\hat{Y}^*_{H-T} = \sum_{i \in s} W_i^* y_i$, where $W_i^* = (\Pi_i p_i)^{-1}$.

[Nargundkar and Joshi (1975), Platek and Grey (1983)].

Since the $p_i's$ are unknown, they must be estimated in some reasonable manner. Naturally the way in which p_i is estimated distinguishes the various adjustment methods.

(a) Adjustment due to Politz and Simmons (1946). This method of adjustment due to Politz and Simmons is based on the idea by Hartley (1946). We have already discussed this method earlier.

(b) Weighting class adjustment: For any probability sample when $W_{ni} = \Pi_{ni}^{-1}, h$ denoting the h^{th} adjustment all, then $\hat{p}_{hi} = \sum_{i=1}^{n_h} W_{hi} \Big/ \sum_{i=1}^{n_h} W_{hi}$.

(c) Post-Stratification adjustment method: the cell-level adjustment for W_{ni} would be $a_b^* = \left(\sum_{n}\sum_{i=1}^{n_b} W_{ni} \Big/ \sum_{i=1}^{n_b} W_{ni}\right)\left(\frac{N_b}{N}\right)$.

[Cohen and Kalsbeex (1981), Bailar etal (1978), Rizvi (1983), Madow (1983), Chapman (1976), Drew and Fuller (1980, 1981)].

Limitations of the method of weighting adjustment:

(a) Computing weights in this manner for survey variables subject to non-response would be time consuming, since adjustment method chosen for each variable must be applied separately;

(b) The analyst doing multivariate analysis involving more than one missing data variable faces the problem of deciding which set of item levels to use, although Little (1988) has proposed a single adjustment based on the regression of the response vector with the vector of variables with no item response.

C2: Imputation:

In the other strategy, gaps remaining after data collection (during and/ or after) are filled by some type of imputation in which a numerical measure replaces the item. In a broad sense, imputation means replacing missing or unusable information with usable data from other sources. These sources can include the same questionnaire (if partial response was obtained), or another questionnaire from the same survey, or external sources, such as another survey or an administrative record.

Let $\underset{\sim}{y} = (y_1, y_2, \mathrm{K}, y_j, \mathrm{K}, y_q)$ corresponds to q questionnaire items. As usual, s denotes a probability sample. Let γ_j be the response set for the variable y_j

i.e., $\gamma_j = \{y_{jK} : K \in s\}$.

The unit response set (i.e., the set of elements that respond to one or more items) is

$$\gamma_u = \gamma_1 \cup \gamma_2 \cup \ldots \cup \gamma_j \cup \ldots \cup \gamma_q,$$

γ_C, the set of elements that provide a complete response vector $\underset{\sim}{y}$ is

$$= \gamma_C \cap \gamma_2 \cap \ldots \cap \gamma_j \cap \ldots \cap \gamma_q.$$

Thus item non-response set is $\gamma_u - \gamma_C$ and unit non-response set $s - \gamma_C$. It has been assumed that both the sets one non-empty.

When there is both unit non-response and item non-response, the estimation of population parameters of interest is more complex. How should the observed data and other information be used? Two options are as follows.

(i) A response set approach. The data associated with j^{th} response set are used to create estimates for the j^{th} variate $j = 1,2,\ldots q$.

(ii) A clean data matrix approach. A completely filled data matrix is created and used to calculate estimates for $j = 1,2,\ldots q$.

A problem with the item-by-item response set approach is that it may lead to impermissible estimates. For example q = 2 and we make estimates of $\sum_u y_{1K}$, $\sum_u y_{1K}^2$ based on γ_1, estimates of $\sum_u y_{1K}$, $\sum_{K \in U} y_{2K}^2$ based on γ_2 and an estimate of the product $\sum y_{1K} y_{2K}$ based on $\gamma_1 \cap \gamma_2$.

If these 5 estimates are used as input for estimating the finite population correlation coefficient between y_1 and y_2, a value may result that falls outside the permissible range [- 1, 1].

The response set approach is thus not without difficulty. To simplify the data handling, survey statisticians usually prefer to work with a complete rectangular data matrix. There are several ways to create such a matrix, which is often called a clean data matrix.

One naïve way to obtain such a matrix is to use the observed y data for $K \in \gamma_C$ only. By treating γ_C as a reduced unit response set, one would then apply the usual techniques for unit non-response. In this method, one disregards observed y-data for elements in $\gamma_u - \gamma_C$. If this set contain few elements, little information is lost. But in other cases, the method could prove wasteful. Instead, imputation is ordinarily used to arrive at a complete matrix.

Imputation implies that an imputed value Z_{jK} is produced for a missing value y_{jK}. The imputed value may be a prediction of the unknown y_{jK}. Auxiliary information may be used to create the value $\underset{\sim}{y}_{jK}$.

It is assumed that values are imputed for the item non-response i.e., an imputation $\tilde{y}_{jK}$ as produced for every missing value y_{jK} such that $K \in \gamma_u - \gamma_C$ and $j = 1, 2K\ q$; this leads to a complete data matrix of dimension $q_{Xn(s)}$.

Different types of Imputation:

A number of imputation techniques have been developed, as discussed by Sande (1982, 1983), Bailar and Bailar (1983), Ford (1983), Kalton and Kasprzyk (1986), Little and Rubin (1987), Little (1988).

Deductive imputation refers to those instances, rare in practice, where a missing value can be filled with a perfect prediction $\tilde{y}_{jK} = y_{jK}$ attained by a logical conclusion. The deduction may be based on responses given to other items on the questionnaire.

Most of the currently used imputation techniques involve the substitution of an 'imperfect' predicted value. Some of the main techniques of this kind are mean imputation, hot deck imputation, cold-deck imputation, regression imputation, and multi imputation.

Overall mean imputation:

This is a simple method that, for item j, assigns the same value, namely, the respondent item mean $\bar{y}_{rj}$ to every missing value y_{jK} in the set $\gamma_u - \gamma_j$, the method may produce a reasonable point estimate of the population total $Y_j = \sum_u y_{jK}$, but is less appealing, if we wish to compute a confidence interval using a standard variance estimator. As is intuitively clear, to replace all missing values for a given item by the respondent mean for that item will give a set of values with less variability than in a sample of equal size consisting entirely of actually observed values. Unless the non-response is negligible or unless a modified variance estimator is used, the method may easily lead to severely understated variance estimators.

Class – mean imputation:

This method works by partitioning the unit response set γ_u into imputation classes such that elements in the same class are considered similar. Auxiliary variables are used for classification. For a given item j, and for all elements k in a given imputation class, missing values are replaced by the respondent mean in that class. There will be some distortion of the natural distribution of y-values, but the distortion is les severe than the overall mean imputation.

Hot – Deck and Cold – Deck imputation:

Improvement on the mean imputation methods is sought by creating a more authentic variability in the imputed values. In hot-deck imputation procedures, missing responses are replaced by values selected from respondents in the

current survey. Cold deck procedures on the other hand, uses imputation based on other sources than the current survey, for example, earlier surveys or historical data.

A number of hot-deck procedures have been proposed, including random overall imputation, random imputation within class, sequential hot deck imputation, hierarchical hot deck imputation and distance function matching.

Random overall imputation:

This method works as follows. For item j a missing value is replaced by an actual observed y_{jK} value taken from a respondent, a donor, randomly drawn from the j response set γ_j. Although the method gives a data set for item j with a close to natural variation, it does not follow that standard techniques can be used straight forwardly, for example, to calculate variance estimates and confidence intervals.

Random imputation within classes:

This is an alternative to the preceding techniques in which suitable classes are formed, similarly as with class mean imputation. For an element in a given class, an imputed value is obtained from a randomly chosen donor in the same class.

Distance function Matching:

This is another hot deck procedure. For item j, a missing value y_{jK} is replaced by the value of respondent classified to be the "nearest" as measured by a distance function defined in terms of known auxiliary variable values.

Regression Imputation:

Unlike the hot-deck imputation, regression imputation uses the estimated relationship between variables. A simple application of this idea is due to Buck (1960) which uses respondent data to fit a regression of a variable for which one or more imputations are needed on other available variables, assumed to have predictive value. The predictors can either be study variables (other high items on the questionnaire) or auxiliary variables. The fitted regression equation is used to produce imputation. For example, for $j = 5$ items with the y variables y_1, y_2, y_3, y_4, y_5, we have

$$\begin{pmatrix} y_{11}, y_{12}, \mathrm{K}\ y_{1K}, \mathrm{K}\ y_{1n} \\ y_{21}, y_{22}, \mathrm{K}\ y_{2K}, \mathrm{K}\ y_{2n} \\ y_{j1}, y_{j2}, \mathrm{K}\ y_{jK}, \mathrm{K}\ y_{jn} \\ y_{51}, y_{52}, \mathrm{K}\ y_{5K}, \mathrm{K}\ y_{5n} \end{pmatrix}$$

Let y_{jK} be the k^{th} value of y_j item. For a certain element $k \in \gamma_u - \gamma_C$, suppose the values y_{1K}, y_{2K} are missing, so that the recorded information for that element reads as follows: Imputations for the blanks are obtained as follows:

$$\begin{pmatrix} -- \\ -- \\ y_{3K} \\ y_{5K} \end{pmatrix}$$

Let $\tilde{y}_1 = f_1(y_3, y_4, y_5)$ be the regression of y_1 on y_3, y_4 and y_5 fitted using data for elements $K \in \gamma_C$

The corresponding estimated regression of y_2 on y_3, y_4 and y_5 is denoted by $\tilde{y}_2 = f_2(y_3, y_4, y_5)$. These two equations and the three recorded values for the element k yield the imputations,

$\tilde{y}_{1K} = \tilde{f}_1(y_{3K}, y_{4K}, y_{5K})$

and $\tilde{y}_{2K} = \tilde{f}_2(y_{3K}, y_{4K}, y_{5K})$

Imputations are calculated analogously for the blanks corresponding to other elements $k \in \gamma_u$.

7. Properly handing of non-response:

The method of handling non-response should be guided by the purposes of data collection efforts and the kinds of non-response. An appropriate method of dealing with non-response should satisfy the following three objectives.

(a) It should adjust estimate for the fact that on measured background, the non-respondents differ from the respondents;

(b) It should expand standard error of estimates to reflect difference between non-respondents and respondents;

(c) It should expose sensitivity of estimates and standard errors to possible difference between non-respondents and respondents on unmeasured background values. The first two objectives are often not satisfactorily addressed, whereas the third objective is usually entirely ignored.

Summary on Dealing with non-responses (Unit):

(a) Before and during data collection, effective measures are taken to reduce the non-response to insignificant levels through some preventive as well as compensatory measures so that any remaining non-response cause little or no harm to the validity of the inferences;

(b) Special, perhaps costly techniques for data collection can be used, for example, randomized response technique .

(c) During data processing stage, some improved method of estimation may be used;

(d) Model assumptions about response mechanism and about relations between variables are used to construct estimators that adjust for a non-response that can not be considered harmless;

(e) Imputation by various techniques.

Summary on Dealing to handle item non-responses:

(a) The basic idea behind many of the above methods used to handle unit non-response can be applied to item non-response as well. Only the notion of non-respondent substitution has been widely adopted from the unit to item non-response level.

(b) Computing item-level adjustment weights might be done to compensate for item-non response.

(c) Little (1988) has proposed a single adjustment based on the regression of the response outcome R_i on the vector of variables with no item response.

Summary on Different Methods of Imputation:

Imputation methods can be classified as

(a) Explicit imputation or (b) Implicit imputation,

both further may be belonging to one of the categories, namely, single valued or multiple valued imputation.

Given the data on the respondents and depending on the mode of analysis, they can further be categorized as one of the followings, (a) Model Based, but non-Bayesian and (b) Model Based, but Bayesian approach.

References

1. Babbie, Earl R. (1973): Survey Research Methods. Belmont, CA, Wadsworth.
2. Backstorm, Charles H. and Gerald Hursh-cesar (1981). Survey Research, 2nd edition, New York, Wiley.
3. Bowley, A.L. (1906): Address to the Economic and Statistics Section of the British Association for the Advancement of Science, York, 1906, J. Roy Statist. SOC. 69, 540-558
4. ----------(1926): Measurement of the Precision attained in Sampling. Bulletin of the International Statistical Institute, 22, 6-62.
5. Bennet, C.M. and Hill, R.E (1964): A comparison of selected personality characteristics of respondents and non-respondents to a mailed Questionnaire. Journal of Educational Research, 58, No. 4 178-180.
6. Birbaum, Z.W. and Monroe G. Sirken (1950): Bias due to non-availability in Sampling Survey. JASA, 45, 98-111.
7. Bailer, Barbara A. (1979): Rotation Sampling Biases and their effects on estimates of changes 43rd session of the International Statistical Institute, Manita.
8. Bergman, L.R. Honve, R. and Rappa, J. (1978): Why do some people refuse to participate interview surveys? Statistik Tidskrift.
9. Brooks, Camilla and Barbara Bailar (1978): An Error Profile Employment as Measured by current Population Survey Statistical Policy Working Paper 3, office Federal Statistical Policy and Standards. U.S. Department of Commerce.

10. Bandyopadhyay, S. Chaudhury, A., Ghosh, J.K. and Maiti, P. (1999): A Draft Proposal for an Enterprise Survey Scheme as a substitute for Economic Census. Indian Statistical Institute, Calcutta.
11. Cochran, W.G. (1977): Sampling Techniques, Wiley Eastern Limited, New Delhi, III edition.
12. Cole, D (1956): Field Work in Sample Surveys of Household Income and Expenditure, Applied Statistics, Volume 5, 49-61.
13. Cobb, J.M., King S., and Chen, E. (1957): Differences between respondents and non-respondents in a morbidity survey involving clinical examination, Journal of Chronic Diseases, 6.
14. Chevry Gabriel (1949): Control of General Census by means of an area sampling method, JASA, 44, 373-379.
15. Chapman, David D. and Rogers Charles, E.(1978): Census of Agriculture- Area Sample design and methodology. Proceedings of the American Statistical Association Section on Survey Research Methods, 141-147.
16. Deming, W (1960): Sampling Design and Business Research, New York, Wiley.
17. ----------(1944): On Errors in Surveys" American Sociological Review, 9, 359-369.
18. ----------(1950): Some Theory of Sampling, John Wiley and Sons, New York.
19. ----------(1953): On a Probability mechanism to attain an Economic Balance between in resultant error and the bias of non-response . JASA 48, 743-772.
20. Dalenius, Tore (1974): The Ends and Means of Total Survey Design; Stockholm, The University of Stockholm.
21. ----------(1957): Sampling in sweden Contribution to the Methods and Theories of Sample Survey Practice, Stockholm, Almquist and Wicksell.
22. ----------(1962): Recent Advances in Sample Survey Theory and Methods, AMS, 33, 325-349.
23. ----------(1977a): Bibliography of non-sampling errors in Surveys. I(A-G), International Statistical Review, 3, 71-89.
24. ----------(1977b): Bibliography of non-sampling errors in Surveys II(A-Q), International Statistical Review, 45, 181-197.

25. ----------(1977c): Bibliography of non-sampling errors in Surveys, III(R-Z), International Statistical Review, 45, 313-317.
26. Dasgupta, A and Mitra, S.N. (1958): A Technical Note on Age Grouping. The National Sample Survey No.12, New Delhi.
27. Dunn, J.P. and Hawkes, R (1966): Comparison of non-respondents and respondents in a Periodic Health Examination Program to a mailed questionnaire, American Journal of Public Health, 56, 230-236.
28. Demaio, T.Y.(1980): Refusals, who where and why? Public Opinion Quarterly 44.
29. Erickson, W.A. (1967): "Optimal Sample Design with non-response", JASA, 62, 63-78.
30. Emrich, Lawrence (1983): "Randomised Response Technique" In William G. Madow and Ingram olkin eds. Incomplete data in Sample Surveys; Volume 2, Theory and Bibliographies, New York, Academic, 73-80.
31. Fellegi, Ivan P. (1963): The Evaluation of the Accuracy of Survey Results Some Canadian Experiences. International Statistical Review, 41, 1-14.
32. ----------(1964): Response Variance and its Estimation, JASA, 59, 1016-1041.
33. Fellegi, Ivan and Sunter, A.B. (1974): Balance between Different Sources of Survey Errors, Some Canadian Experiences, Sankhya, 36, Series C), 119-142.
34. Ferver, Rebert (1966): Items non-response in a consumer survey, Public Opinion Quarterly, 12, 669-676.
35. Ford, Barry L. (1976): Missing Data Procedures, A Comparative Study, American Statistical Association, Proceedings of the Social Statistics Section 1976, Pt. 1, 326-329.
36. Ghosh, J.K. and Maiti, Pulakesh (2003): The Indian Statistical System at cross roads an appraisal of Past, Present and Future, To be presented at the IMS meet during 2-3 January – 2004.
37. Ghosh, A (1953): Accuracy of Family Budget Data with reference to period of re-call, Calcutta Statistical Association Bulletin, 5, 16-23.
38. Gower, A.R. (1979): Characteristics of non-respondents in the Labour Force Survey, Statistics Canada.

39. Groves, Robert, M. and Kahn Robert Louis (1979): Surveys by Telephone, A national comparison with personal interview, New York; Academic.

40. Gray, P. and Gee, F.E.N. (1972): A Quality check on the 1966 ten percent sample census of England Wales, office of the population census and surveys, London.

41. Ghosh, J.K., Maiti, P. Mukhopadhyay, A.C., Pal, M.P (1977): Stochastic Modeling and Forecasting of Discovery, Reserve and Production of Hydrocarbon-with an application, Sankhya, Series B, 59, pt. 3, 288-312.

42. Hansen, M.H., Madow William G., and Tepping B.J. (1983): An Evaluation of Model dependent and Probability Sampling inference in Sample Surveys, JASA, 78, 776-807.

43. Hanse, M.H., Hurwitz William N. (1946): The Problem of non-response in Sample Survey, JASA, 41, 516-529.

44. ---------and Nisselson, H., Steinberg, J. (1955): The redesign of the current population survey, JASA, 50, 701-719.

45. ----------Jubine, Tomas B. (1963): The use of imperfect lists for Probability Sampling at U.S. Bureau of Census, Bulletin of the International Statistical Institute, 40(1), 497-517.

46. --------and Pritzker, Lenon (1964): The Estimation and interpretation of Gross differences and the simple response variance. In C.R. Rao with D.B. Lahiri, K-P.

47. Nair, P. Pant and S.S. Shrikhande eds. Contributions to Statistics Presented to Professor P.C. Mahalanobis on the occasion of his 70th birth day Oxford, England, Pergaman, Calcutta Statistical Publishing Society, 111-136.

48. ----------and Bershad, Max A. (1961): Measurement errors in censuses and surveys, Bulletin of the International Statistical Institute, 38, 359-374.

49. ----------Marks, Elis Mauldin, Parker W. (1951): Response Errors in Surveys, JASA, 46, 147-190.

50. ----------(1976): Some Important Events in the Historical Development of Sample Surveys in Donald Bruce Owen ed., on the History of Statistics and Probability, Statistics Text Books and Monographs, Volume 17, New York Dekker, 73-102.

51. Hurscgberg, David Frederick, J. Scheuren and Yuskavage Robert (1977): the impact on Personal and Family income of adjusting the current population survey for under

coverage, Proceedings of the Social Statistics Section, American Statistical Association, 70-80.

52. Hubback, J.A. (1927): Sampling for rice yields in Bihar and Orissa, Imp. Agr. Res. Inst. Bulletin, Pusha (reprinted in Sankhya (1946), 7, 282-294)
53. Halden, J.B.S. (1957): The Syadvada System of Prediction, Sankhay 18, 195-2000.
54. Hacking, J. (1965): Lobgic of Statistical Inference Cambridge University Press.
55. Hoinville, Gerald and Robert Joell (1978): Survey Research Practic , London, Heinemann.
56. Jessen, Raymund J. (1978): Statistical Survey Techniques, New York, Wiley.
57. Kiawer, A. (1895): Observations et experiences concernant des denombrements representatives, Bull. Int. Statist. Inst. 9, 176-183.
58. Kruskel, William and Frederick Mosteller (1980): Representative Sampling, IV, the History of the concept in Statistics, 1895-1939, International Statistical Review, 48, 169-195.
59. Kish, L. (1965): Survey Sampling Wiley and Sons, New York.
60. ------- and Hess I. (1958): on non-coverage of sampling dwellings, JASA, 54, 509-524.
61. Kalton, Graham and Daniel Kasprzyk (1982): Imputing of missing survey Response, American Statistical Association 1982, Proceedings of the Section on Survey Research Methods, 22-31.
62. Koop, J.C. (1974): Notes for a unified theory of estimation for sample surveys taking into account response errors, Metrika, 21, 19-39.
63. Kalton, Graham (1983): Compensating for missing survey data Research Report Series, Ann. Arbor M1, Institute for Social Research, University of Michigan.
64. Kendal, Maurice George and William R. Buckland (1960): A Dictionary of Statistical Terms, 2nd edition, London, Oliver and Boyd.
65. Lubin, B. Levitt, E. and Zuckerman, M. (1962): Some personality differences between respondents and non-respondents in a survey questionnaire, Journal of Consulting Psychology, 26-192.
66. Lundberg, G.A. and Larsen, O.A. (1949): Characteristics of Hard-to-reach individuals in field surveys, Public Opinion Quarterly, 13, 487-494.

67. Lyberg, L. and Rapaport, E. (1979): Unpublished non-response problems at the national central Bureau of Statistics, Sweden.

68. Little, Roderick J.A. (1982): Models for non-response in Sample Surveys, JASA, 77, 237-250.

69. ----------(1983): Super Population models for non-response, Part IV. In William G. Madow and Ingram Olkin eds. Incomplete data in Sample Surveys, Volume 2, Theory and Bibliographies, New York, Academic, 337-413.

70. Lessler, J.T. (1974): A double sampling scheme model for eliminating measurement process bias and estimating measurement errors in surveys, Institute of Statistics Mimeo Series No. 949, University of North Carolina, New Chapel.

71. ----------(1980): Errors associated with the frames, Proceedings of the American Statistical Association Section on Survey Research Methods, 125-130.

72. Madow, W.G. Nisselson, Harold and Olkin, Ingram (1983): Incomplete data on Sample Survey, Volume 1, Report and Case studies; New York, Academic.

73. Mc. Neil, John M. (1981): Factors affecting the 1980 census content and the effort to develop a post census disability survey. Presented at the annual meeting of the American Public Health Association.

74. Mahalanobis, P.C. and Lahiri, D.B. (1961): Analysis of errors in censuses and surveys, Bulletin of the International Statistical Institute, 38(2), 359-374.

75. ----------and Sen, S.B. (1954): On some aspects of the Indian National Sample Survey, Bulletin of the International Statistical Institute, 34, pt. 2.

76. Mahalanobis, P.C.(1944): On Large Scale Sample Surveys, Philosophical Transactions of Royal Society, 231-(B), 329-451.

77. ----------(1946): Recent Experiments in Statistical Sampling in the Indian Statistical Institute.

78. ----------(1941): A Sample Survey of the Acre-age under jute in Bengal, 4, 511-30.

79. ----------(1954): The Foundations of Statistics, Dialectica, 8, 95-111 (reprinted in Sankhya 18, 183-194)

80. Maiti, P (1983): unpublished Ph.D. Thesis entitled Some Contributions to the Sampling Theory using auxiliary information" submitted to the Indian Statistical Institute, Calcutta.

81. ----------, Pal, M. and Sinha B.K. (1992): Estimating unknown Dimensions of a Binary matrix with application to the estimation of the size of a mobile population. Statistics and probability, 220-233.

82. Moser, Clays Adolf and Graham Kalton (1972): Survey methods in Social investigation , 2nd edition, New York, Basic Books.

83. Mooney, H. (1962): On Mahalanobis' contributions to the development of sample survey theory and method in C.R. Rao etal (eds) contributions of statistics, Pergamon Press.

84. ---------- (1967): Sampling Theory and Methods, Statistical Publishing Society, Calcutta.

85. Neyman Jerzy (1934): On the two different aspects of the representative method, the method of stratified sampling and the method of purposive selection, J. Roy. Statist. SOC. 97, 5589625.

86. Neter, J. and Waksberg, J. (1965): Response errors in collection of Expenditures data by household interview. An Experimental Study Technical Report No. 11 U.S. Bureau of the Census.

87. Newman, S. (1962): Difference between early and late respondents in a mailed survey, Journal of Advertising Research, volume 2, 37-39.

88. Ognibene, P. Traits affecting questionnaire response, Journal of Advertising Research Volume 10, 18-20.

89. Pan, J.S. (1951): Social Characteristics of respondents and non-respondents in a questionnaire study of later maturity, Journal of Applied Psychology, 35, 780-781.

90. Politz, A.N. and Simmons, W.R. (1949): An attempt to get Not-at Homes into the sample without call-backs, JASA, 44, 9-31.

91. Platek, R. (1977): Some factors affecting non-response, Survey Methodology, 3.

92. Plan, V.T. (1978): A Critical appraisal of household surveys in Malaysia Multipurpose household survey in developing Countries, Development Centre, OECD, Paris.

93. Reuss, C.F. (1943): Differences between persons responding and not responding to mail questionnaires, American Sociological Review, 8, 433-438.

94. Rao, V.R. and Sastry, N.S. (1975): Evolution of a total survey design, The Indian Experience, Invited paper presented to the International Association of Survey Statisticians Warsaw.

Report of Research Projects

(1975-76): Cost Benefit Analysis of Rural Electrification, Project Leader Professor J. Roy, Computer Science Unit, ISI, Calcutta.

(1977-78): Calcutta Urban Poverty Survey, Project Leader Professor J. Roy, Computer Science Unit, ISI, Calcutta.

(1988-89): A Survey on Domestic Tourists in Orissa, Project Co-ordinator, P. Maiti, ISI, Calcutta.

(1994-95): An Enquiry into the Quality of Life in five communities in selected districts of Rural West Bengal, Project Co-ordinator, P. Maiti, ISI, Calcutta.

(1995): Community attitudes and Preferences pertaining to cemetery and cremated related issues in the East Rand in the Republic of South Africa, CENSIAT, HSRC, Pretoria, South Africa, Principal Statistician – P. Maiti.

(1995): Survey of family and Community life in the Selected Communities of the cape Peninsula of the Republic of South Africa, CENSTAT< HSRC, Principal Statistician – P. Maiti.

(1995): the Socio-economic demographic and cultural pattern of the female labour force participation in the North West and the Cape; CENSTAT, HSRC South Africa, Principal Statistician – P. Maiti.

(1996): Stanza-Bopape Project; CENSTAT, HSRC, South Africa, Principal Statistician – P. Maiti.

(1998): Mid. Term Review of IPP-VIII in Calcutta Metropolitan Area, ISI, Calcutta, Survey Statistician – P. Maiti.

(2001, August): National Statistics Commission , Government of India.

95. Roshwalb, Alan (1982): Respondent Selection Procedures within Households, American Statistical Association 1982 Proceedings of the section on Survey Research Methods, 93-98.

96. Rubin Donald B. (1983): Conceptual issues in the presence of non-responses, In William G. Madow and Ingram Olkin eds. Incomplete data in Sample Surveys 2, Theory and Bibliographies, New York, Academic, 123-142.

97. ----------(1977): formalizing Subjective notions about the effect of non-respondents in Sample Surveys", JASA, 72, 538-543.

98. ----------(1978): Multiple imputations in Sample Surveys – A Phenomenological Bayesian Approach to non-response, American Statistical Association 1978 proceedings of the Section on Survey Research Methods, 20-28.

99. ----------(1987): Multi imputation for non-response in Surveys, New York, Wiley.

100. Rizvi, M. Haseeb (1983): Hot-Deck Procedures Imputation in William G. Madow and Ingram Olkin eds., Incomplete Data in Sample surveys, 3, Proceedings of the symposium, New York, Academic, 351-352.

101. Stephen, Frederick F. (1948): History of the uses of Modern Sampling Procedures, JASA, 43, 12-39.

102. Smith, T.M.F. (1976): The Foundations of Survey Sampling, A Review, JRSS, 139A, 183-195.

103. Sarndal, C.E., Swensson, B. and Wretman, J. (1992): Model Associated Survey Sampling , Springer Verlag, New York, Inc.

104. Scott Christopher (1973): Experiments on recall error in African Budget Surveys, paper presented to the International Association of Survey Statisticians, Viena.

105. Shah, Nasra M. (1981): Data from Tables used in the paper presented at Weekly Seminar of East West Population Institute, October 28, Honolulu.

106. Sudman, Seymour (1976): Applied Sampling, New York, Academic.

107. Suchman, Edward A. (1962): An analysis of bias in Survey Research, Public Opinion Quarterly, 26, 102-111.

108. Scheaffer, Richard, L. Mendenhall, William and Ott lyan (1979): Elementary Survey Sampling, 2nd edition, North Scituate, MA: Duxbury Press.

109. Szameitat, Kleus and Schaffer, Karl August (1963): Imperfect Frames in Statistics and the consequences for their use in sampling, Bulletin of the International Statistical Institute, 40, 517-544.

110. Singh, Bahadur, Sedransk, Joseph (1978): A two phase sampling design for estimating the finite population mean when there is non-response, In N. Krishnan Namboodiri ed. Survey Sampling and measurement, New York, Academic, 143-155.

111. Thompson, Ib and Siring, E. (unpublished): On the causes and effect of non-response, Norwegian Experiences, Central Bureau of Statistics, Norway.

112. Tuygan, Kuthan and Cavador, Tevfik (1975): Comparison of self and presale responses related to children of ever married women. In laboratories for population Statistics Scientific Report Series No. 17, 22-28.

113. Turner, Anthony G., Waltmen, Henry. F, Fay Robert and Carlson Beverly (1977): Sample Survey Design in developing Countries – three illustrations of methodology, Bulletin of the International Statistical Institute.

114. U.S. Bureau of the Census (1974): Standards for the discussion and presentation of errors in data, Technical Report No. 32.

115. ----------(1976): An overview of population and housing census evaluation programmes conducted at the U.S. Bureau of Census, Census Advisory Committee of the American Statistical Association.

116. Verma, Vijay (1980): Sampling for national fertility surveys, World fertility survey conference, London.

117. Warwick, Donald P. and Chartes A. Lininger (1975): The Sample Survey, Theory and Practice, New York, Mc. Graw Hill.

118. Woltman, Henry and Bushery, John (1975): A panel bias study to the national crime survey. Proceedings of the Social Statistics Section. American Statistical Association.

119. William, W.H. and Mallows, C.L. (1970): Systematic biases in panel surveys JASA, 65, 1338-1349.

120. Warner, Stanley L. (1965): Randomised Response, A Survey Technique for eliminating Evasive answer bias, JASA, 60, 63-69.

121. Zarkovich. S.S. (1966): Quality of Statistical data, Rome; FAO of the United Nations.

Non-Response in Surveys: Reasons, Consequences And Prescriptions for Their Control

Arijit Chaudhuri
Applied Statistics Unit,
Indian Statistical Institute,
Kolkata

A well-known disquieting feature in Censuses and Sample Surveys is the phenomenon of "Non-Responses". This inevitably induces 'Errors' in surveys. Suppose our purpose is to ascertain the socio-economic conditions currently prevailing in a specific community. A possible way to achieve this is to clearly specify the individuals we intend to cover and then address a suitably designed questionnaire either to each of them or to a suitably chosen sample out of them. Usually the parameters drawing our attention are totals of one or more variables of interest taking values on these units or of some simple functions of these totals. In case of a Census if values turn out missing for some of the units, the parameter values cannot be accurately ascertained. In case of a sample survey the sampling design prescribes certain suitably weighted sums of values of the sampled units to provide appropriate estimates for the parameters defined by aggregation over the unit-values. In case variate-values turn out missing for some of the units sampled then the

originally prescribed weighted sums of the values of the available sample units fail to be correct estimates for the intended totals which are the parameters.

Many reasons can be ascribed for such 'misses' or 'non-responses'. knocking at the door of a person intended to be covered may not find him/her 'At Home' to respond. A person to be covered may not allow an interviewer any access to himself/herself. A person interviewed may not agree to answer any question at all or some of the questions at least addressed to him/her. This may be because he/she may not give out truths about himself/herself because of personal dislikes. The interviewer may not be persuasive enough to elicit a response needed. On providing suitable inducements a 'repeat' effort may be more successful. In case of surveys other than on human behaviour at least directly, for example, a crop survey, because of inhospitable location of the crop field it may be difficult to gather data on crop production leading to misses.

The discrepancy between an ascertained value of a parameter and the true parametric value in case of a census is called 'Bias'. Likewise the expected value of an estimator minus the parameter it seeks to estimate is called the bias of the estimator.

There are essentially two distinct approaches to control bias in an estimator arising out of ;Non-Response'. One is 'Weighting Adjustment' to cover the case when there is non-response entirely for a selected unit. In case there is only partial non-response in respect of only a few items in a questionnaire while response is available on other items from a respondent the method used to tackle is called an "imputation Technique".

Suggested Further Reading Materials :

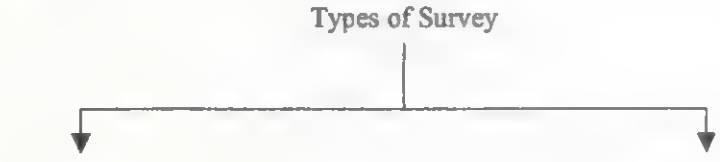
1. Madow, W.G. & Olkin, I. Eds (1983). incomplete data in sample surveys. Vol I Academic Press, NY.
2. ___ & Olkin,I. Eds (1983) Do Vol III, Do
3. __, ___ & Rubin, D.B. Eds (1983), Do Vol II, Do, Do
4. Rubin, D.B. (1987). Multiple Imputation for Non-Response in surveys. Do, Do.

MANAGING DISSIMILARITY IN SCALING PROPERTIES OF SCHEDULE

D. Dutta Roy
Psychology Research Unit
Indian Statistical Institute
203, B.T. Road
E-mail: ddroy@isical.ac.in

Analysis of survey data falls on two classes – (a) obtaining descriptive information about estimates of population characteristics and (b) obtaining information about relationship among different population characteristics. Former is applied in the census type and the later is in the investigational type of survey research. In the census type of survey, estimates of the characteristics of the whole population and possibly of various previously defined subdivisions of it are required. Therefore analysis of data mainly reflects descriptive information about the population. These descriptive information form the basis of administrative action, either directly or after incorporation with information from other sources. On the other hand, investigational type of survey is more concerned with the study of relationship between different variates, with contrasts between different domains. In such surveys, estimates appertaining to the whole population are usually of relatively minor interest. It would be often erroneous to make conclusion about relationship among different statements or variates or domains based upon census survey data. Therefore, investigational survey data analysis is important.

The critical analysis of the results of an investigational survey is much more difficult task than is the calculation of estimates and their errors in a survey of the census type. Dissimilarity in scaling properties of different items in the schedule make analysis more critical.



(Estimation of population characteristics and errors or descriptive information)

(Analyzing relations among different characteristics of sample and making inference or relational information.)

In case of investigational survey, therefore attention should be paid to scaling similarity among different domains or variates of population characteristics or of human behaviour. Following steps may be considered in design of good schedule:

1. Framing hypotheses of the survey ;
2. Identify domain and sub domain of explanatory, intervening and dependent variables ;
3. Scaling the domain and sub domain and operationally define each domain so that they can be assessed objectively ;
4. Develop statements for assessing each domain or sub domain ;
5. Classify the response categories related to each statement ;
6. Scaling the response categories following prior studies or theory ;
7. Prepare the Table of cross tab analysis or graphical distribution to show the relationship among different response categories ;
8. If Tables and graphical representation are theoretically meaningful, then accept them otherwise start thinking from step 2.

Framing hypotheses of the survey

Investigational survey aims at testing assumed model and the result provides map of relationship among set of variables, therefore in framing hypotheses, one should pay more attention to the model development. Again, in model building, one can think of validity of schedule by correlating set of statements of specific domain.

Identify domain and sub domain of explanatory, intervening and dependent variables.

Each variable accounted for the survey possesses set of population characteristics or behaviour domain. Domain may be uni or multi dimensional. Researcher initially should think of the nature of domain based

on the objective of the study, prior studies , group discussion with target people and on time limit.

Scaling the domain and sub domain and operationally defining each domain so that they can be assessed objectively.

In operational definition of each domain, attention should be paid to the measurement scales (e.g. to what extent) rather on mere description.

Develop statements for assessing each domain or sub domain.

Statements should be in line with the attributes of each domain. Keep It Simple and Specific (KISS). Simple means unambiguous and specific refers to the culture of target group.

Classify the response categories related to each statement

Response categories should be classified based on theory. Number of categories depends upon the mental set up of target group so that target group will not be confused in making judgement.

Scaling the response categories following prior studies or theory.

Considering scale properties, one may classify the response categories in terms of four scales as nominal, ordinal, interval and ratio. Nominal scales merely classify without indicating order, distance or unique origin. Ordinal scales indicate magnitude of relationships of more than or less than but indicate no distance or unique origin. Interval scales have both order and distance values but no unique origin. Ratio scales possess all three features. Selection of scales depends on the objective or hypotheses of study.

Prepare the Table of cross tab analysis or graphical distribution to show the relationship among different response categories.

Cross tabulations generally allow us to identify relationships between the cross tabulated variables. It provides insight about number and direction of response categories. Similarly graphical distribution of relationship (positive, negative or zero) provides above insight. Cross tabulation is specially useful in case of categorical and the later one is in case of non categorical variables.

The above discussion is specially useful during schedule design or pre survey stage. But there are some conditions when researchers are confronted with data having dissimilar scaling properties or having multiple scale combinations, during that condition they can think of following strategies in analysis of data.

Table 1
Managing dissimilarity in scaling properties of schedule

Groups	Strategies	Statistics	Comment
Nominal and Interval	Convert interval scale into nominal based on cutting point	Chi-square	Loss of minor deviation effect of interval scale.
	If no conversion	Biserial or Point biserial correlation	Applicable only for two sets of variables.
Nominal and Ordinal	Convert ordinal scale into nominal based on cut-point.	Chi-square	Loss of minor deviation effect of ordinal scale.
Nominal and Ratio	Convert ratio scale into nominal based on cut point.	Chi-square	Loss of minor deviation effect of ratio scale.
Interval and Ordinal	Converting ordinal scale value into score or interval scale	Correlation	Loss of scaling properties of order.
Interval and ratio	No conversion is required	Correlation	No loss of statistical properties

CASE STUDIES

Relationship between writing motivation and school categories (Interval VS Nominal data)

Table 2
Frequency distributions of Writing motivation variable scoring categories across school types

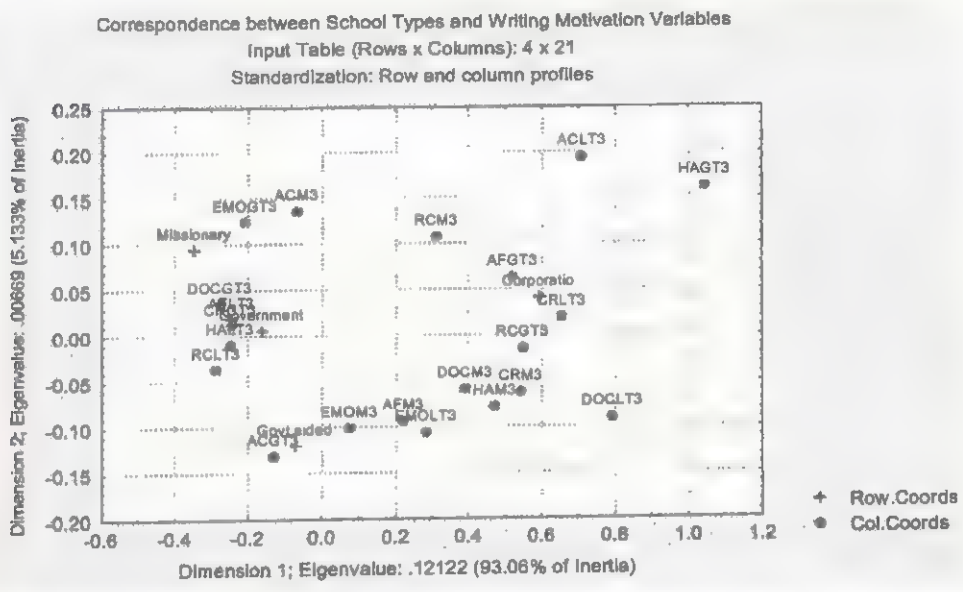
	Government (n=205)	Govt. aided (n=252)	Corporation (n=212)	Missionary (n=215)	Chi square(df=6)	P Value
DOCGT3	148	171	78	192		
DOCLT3	13	34	65	5		
DOCM3	44	47	69	18	152.72	0.0001
EMOGT3	100	99	68	133		
EMOLT3	52	68	77	29		
EMOM3	53	85	67	53	50.35	0.0001
CRGT3	167	181	87	188		
CRLT3	18	25	54	9		
CRM3	20	46	71	18	132.37	0.0001
HAGT3	11	16	68	4		
HALT3	174	198	90	196		
HAM3	20	38	54	15	181.06	0.0001
AFGT3	32	33	70	18		
AFLT3	132	152	74	161		
AFM3	41	67	68	36	86.91	0.0001

ACGT3	114	166	82	114		
ACLT3	24	16	62	11		
ACM3	67	70	68	90	86.29	0.0001
RCGT3	27	43	76	18		
RCLT3	145	164	62	156		
RCM3	33	45	74	41	116.31	0.0001

Chi-square (60)=806.007, p<0.0001

Figure 1: Plotting correspondence between writing motivation and school types

Case study 2



Plotting map of mental health distribution across different months in the Antarctica expedition.

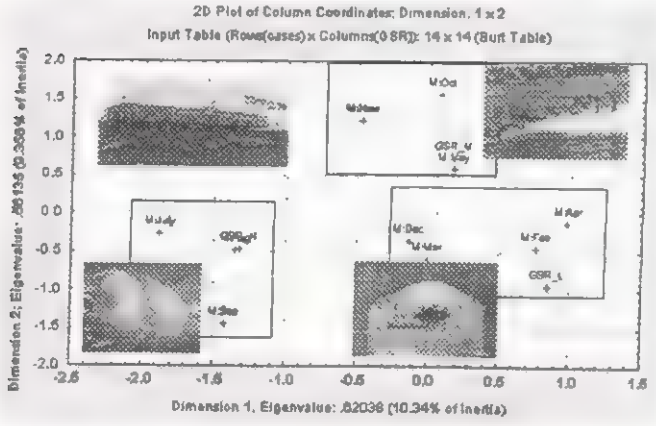
Data were collected from 11 scientists and 8 logistic personnel across 11 months (from Feb to Dec.) using GSR. Table 3 shows Burt Table of few input data. GSR scores were classified into 3 groups based on quartile. In the Burt Table, 1,2, and 3 codes were used wherein 1=low stress (score less or equal to 434) , 2=Moderate stress, (score is within the range from 435 to 495) and 3= High stress (score greater than 495).

Table 3
Burt Table of month and GSR data

SL	MONTH	GSR CODE
1	2	2
2	2	2
3	2	2
4	3	1
5	3	1
6	3	1
7	4	2
8	4	2
9	4	2

Significant chi-square value ($\chi 2$ (169) = 4639.79, $p<0.00$) suggests variation of stress across months. Correspondence map (Figure 2) shows high stress level during the months of July, August and September. Moderate level of stress was noted in the months of Feb, Mar, April and December. And lower level of stress was found in the months of May, June and October. One must be cautioned in interpreting the correspondence map as

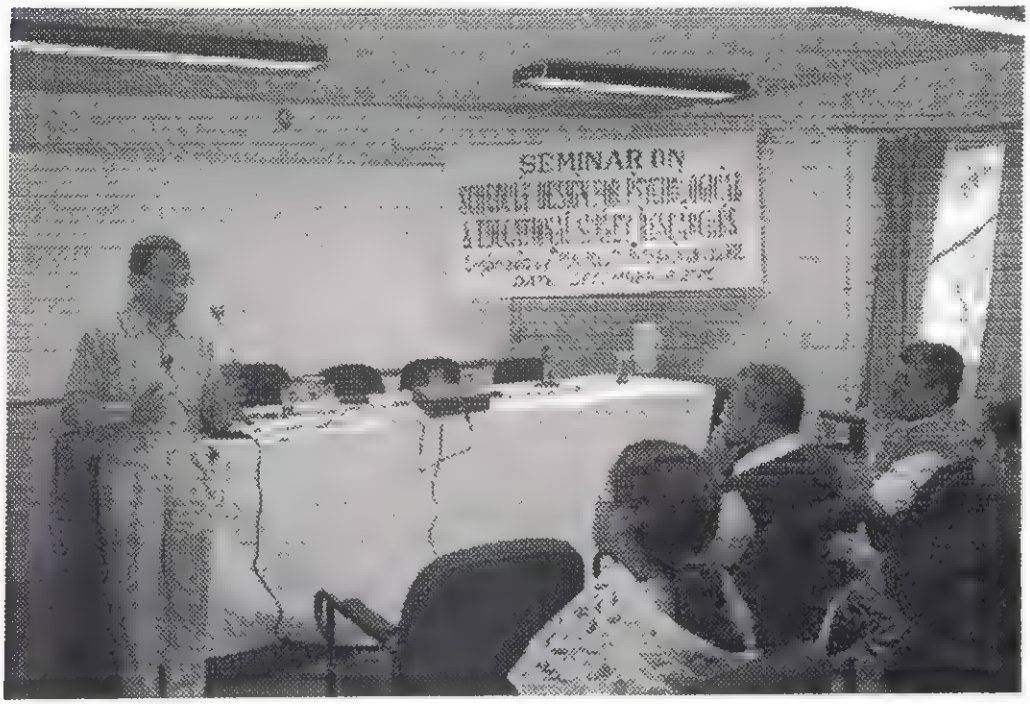
Figure 2
Correspondence map of GSR score and Months



PILOT TESTING OF SCHEDULES

S.P.Mukherjee

Centenary Professor, Department of Statistics,
Calcutta University
Kolkata



Schedules and questionnaires are commonly used instruments to collect items of relevant information from some identified individuals. Responses to various items in the schedule provided by some or all of the identified individuals are checked, cleaned and subsequently analysed to reveal facts and figures bearing on the objective (s) of the study. In large-scale surveys meant to cover a large number of potential respondents as also in small-scale studies meant to provide an in-depth analysis of a phenomenon in which we are interested, a pilot survey is often carried and findings there from are taken into account to finalise various facets of the ultimate study.

One important component of a pilot survey is a pilot- or a pre-testing of the schedule or questionnaire to be canvassed. This testing may be needed for various purposes. The following gives an indication.

1. This will provide a provisional estimate of the time to be taken (and hence the cost involved) in canvassing a single schedule so as to yield an estimate ot the total cost for canvassing a given number (as indicated in the sample size) of respondents. Alternatively, given the total resource available for the study, this estimate may help us in determining the sample size. Of course, there are other criteria to taken into account while determining the sample size.

2. Pilot testing will tell us whether the different items included in the schedule are necessary and sufficient to bring out the study objective (s). talking of sufficiency, one may think of some apparently redundant items which may be required to check internal consistency among responses to related items.

3. A very important purpose served by pilot testing it to find out if the questions or statements in the schedule are unequivocally understood by the potential respondents or not. The language used must be appropriate to the group being canvassed. Depending on the feedback from interviewers engaged in pilot testing, the language and the over-all presentation of the schedule may have to be modified.
4. Findings from a pilot testing exercise may reveal that even if the individuals approached could clearly comprehend the questions or statements, they are not well-informed about the underlying issue (s) and hence cannot provide responses that can throw light on such issues.

5. A pilot test may also bring to light the fact that some of the items in the schedule relate to delicate or sensitive matters or matters which are perceived to encroach on the interviewee's privacy .In such cases, if the concerned are responses are really needed in the study, one can go for Randomised Response Technique.

6. The points mentioned in the previous two sections will be of immense use in finalizing the sampling frame for our study as well as the sampling design to be adopted.

7. Pilot testing may also reveal the need for giving leads to interviewees by investigators and, designed properly, can also reveal by way of differences the impact of such leads on the responses. Eventually, this will help in training the investigators about leads to be given by them-their nature and extent. Pilot testing may serve other useful purposes also. It must be remembered that a careful pilot testing exercise involves a bit of research and hence should be carried out by competent supervisors, if not by the project seniors themselves. The design for a pilot testing will depend on the purposes it is expected to serve. IN fact, canvassing a schedule during the pilot testing phase will surely involve longer time than what will be subsequently required by trained investigators in the final study.

8. Data collected through a pilot test should be analysed primarily to reveal differences among interviewee groups, among different response patterns corresponding to different types of leads given, etc.

ROLE OF PILOT SURVEY IN SOCIAL SCIENCE RESEARCH METHOD

Professor Prafulla Chakrabarti
Director of Research, SERI
Mohana, 5, New Raipur,
Kolkata.

Introduction

Perhaps not many social researchers will deny the fact that no organized schedule or questionnaire can be constructed without a priori knowledge gathered from Pilot Studies and Pretests. No amount of thinking, no matter how logical the mind or brilliant the insight is, a schedule or questionnaire cannot be standardized without careful empirical checking. Unfortunately, many a students of social science discipline do not have patients to follow these basic principles because of lack of proper training and guidance in field- work methodology or they are too impatient to spend time on pilot study and pre-testing in schedule construction. It may be because of the notion that construction of schedule is not a very difficult task and anybody can make a schedule if he desires to investigate some phenomena. It may also be the fact that owing to their over enthusiasm, considering that refinement of tool is an act of unnecessary wastage of

time and energy. More often, they prepare "so called schedule" carelessly and collect items of information from the field that are not only irrelevant but their authenticity begs question. After collection of a huge mass of haphazard information which a researcher labeled as data (?) they usually approach their teachers to know what shall they do with them, how would they relate these with their objectives and so on.

With my more than five decades of involvement in social science research and basic training in methodology from no less than the masters like T.C. Das, N.K. Bose, Irawati Karve, Ramkrishna Mukherjee, Nikhilesh Bhattacharya and so on, I shall try to discuss in what ways Pilot study and Pretests help researcher in framing a schedule. In other words, I would address myself to the questions:

1.What is Pilot Study?
2.Why should one make Pilot Study?,
3.How it helps researcher to improve their study design?,
4. Is it necessary for all specialties in social science discipline? and ,
5.What is the relationship between Pilot Study and Pretests?

1. What is Pilot study?

The procedure by which a researcher formulates items in areas where the literature is scanty (e.g., suppose one researcher wants to frame a schedule on elder abuse – a rare but increasingly becoming social problem in the field of social gerontology, but hardly has any authentic material) and the, manner by which he selects items for the final schedule is called Pilot Study and Pretests. Suffice it to say, these are standard practice with professional survey bodies and are widely used in research surveys.

2. Why should one make Pilot Study?

Besides the few remarks made before, the Pilot survey is of utmost importance for framing a sound schedule for a number of reasons :

These are as follows:

(i) It checks the adequacy of the sampling frame from which it is proposed to select the same.

Examples :

a) A reader of a reputed university planned to estimate the rate of drop-outs in primary school level students for which he had consulted the school registers with a view to making a sampling frame. He did not try before hand whether a student was on leave for treatment, or got failed, for other contingent reasons and so on.

b) Or take another example. A researcher might be planning to use the pay-roll of workers in a factory as the basis for drawing a sample. He was not aware another some workers on leave whom the survey was to include or cards may be temporarily removed from it when required for some other purpose, whether or not such defects could have been overcome is another matter, what is vital is to be aware of them before starting the survey.

(ii) It is well known that the decision to determine sample size requires some a priori knowledge of the variability of the population. The Pilot study provides valuable supporting evidence.

(iii) The probable numbers of refusals and non-contacts can roughly be estimated from the pilot survey and partially from pretests. The effectiveness of various ways of reducing non-response can be compared. As a result, one data-collecting method may be chosen in preference to another, some questions may be excluded, the timing of interviews may be changed and so on. For example, one may debate whether to collect data from a widely dispersed population by mail or interview. The former is cheaper but will it achieve an adequate response? The answer is NO. The study of Data Inventory in Social Sciences a UNESCO funded project undertaken by the SRU is an eye opener.

(iv) From pilot survey and pretests we can know whether the interviewers are doing an efficient job, whether excessive strain is being placed on them or on the respondents.

(v) The adequacy of the questionnaire can be judged from the pilot survey. This is probably the most valuable function of pilot survey. The pilot survey and pretests offer a way of trying the questionnaire with the kind of interviewers and respondents in the main survey. Other points, the case of handling the questionnaire in the field, the efficacy of its layout, the clarity of the definition etc can be checked. Is the wording simple. Clear, unambiguous, free from jargons? Answers can be had from pilot survey.

Ilustrations from the study of Social attitude towards Air Pollution sponsored by DST and carried out by ISI, How do the poor Survive ? by ISI may be offered.

(vi) The efficiency of the instructions and briefing of Interviewers can be judged from the scrutiny of the completed questionnaire.

For example, interviewers did not put a ring round the "Not Applicable" code when a question did not apply; they omitted their own identity number on the questionnaire etc.

Illustration from The Study of West Bengal Family Structure by ISI can be made.

(vii) Without the Pilot Survey it is often hard to decide on the alternative answers to be allowed for in the coding. One may wish to ask, "what are the furniture do you have in your home" and to print all or as many as answers one may think of. In the pilot survey, this can be asked as an 'open' question.

(viii) If it appears that the survey will take too long or too expensive, the pilot survey can be valuable in suggesting where economics can be made.

(ix) The pilot survey nearly always results in important improvements to the questionnaire and a generally increases in the efficiency of the enquiry. It is the last safe-guard against the possibility that the main survey may be ineffective.

Examples : Study of the Sarak Culture and Society, Value Systems and Social Change, Impact of Culture on Housing etc.

(x) The size and design of the pilot survey is a matter of convenience time and money. It should be large enough to fulfill the above functions and the sample should ideally be of a comparable structure to that of the main survey.

3. How does Pilot study improve study design?

We have seen that the pilot survey can help to guide the choice between alternative methods of collecting data (Study of Elder Abuse – can be illustrated), ordering the questions, wording and so forth. It should be designed therefore so as to ensure a strict testing of these alternatives. If two forms of a question are to be compared, each should be tried out on an equivalent sample of respondents; otherwise the difference in effectiveness of the two questions would be mixed up with difference between the samples themselves.

It many types of comparisons are to be made simultaneously – between interviewers, questions, non-response methods instructions and so on this calls for strict methods of experimental design.

4. Is pilot survey sine-qua non in every research belonging to all social science specialties ?

Before addressing the question let me confess that I do not believe demography, anthropology, sociology, psychology, education, social work, history, geography (human) etc are distinct social science disciplines. Because each explore reality in its own way and thus appraise the social reality only partially and thus each one at best can be regarded as separate specialty under the rubric of social science discipline. However, it should be made clear that those specialties which rely heavily on empirical research, and depends on the instrument of schedule and questionnaire for collecting data, pilot survey is extremely important. In some methods of data

collection in anthropological research like participant observation, Fussed Group interview, Genealogical, Rapid Rural Appraisal, Participatory Research Method – pilot survey is not so important. For research on history, pilot survey is not required, and it is also not so much important for the armchair social scientists who are quite often dependent on secondary data.

5. Relationship between Pilot Survey and Pre-tests.

If pilot survey is a safe-guard for any loop-holes in constructing a schedule or questionnaire, Pretests are a dress rehearsal prior to the implementation of the main show of data collection. After completion of the pilot survey, the researcher should be ready to set up a pre-test procedure. This is a much more formal step than a pilot study. It entails that every part of the procedure must be laid out exactly as the final study will be carried out. The interviewing instructions is to be used, the cover letter or instructions should be put in final form.

Illustration : Study of Drug Addicted Youths in Calcutta.

A good researcher will actually tabulate the data from this pretest in order to see what weakness are present. This will include the proportion of do not know' answers for difficult, ambiguous and poorly warded questions, the proportion of respondents who refuse to be interviewed etc. In short, the pretest accepts the fact that no amount of intuition, native talent, or systematic thought will substitute for the careful recording, tabulating and analysis of the research fact. These facts must be obtained before the final investment of much time, money and energy in a full-scale research project.

Conclusion

The purpose of research is to unravel the truth. For social science research, the purpose is more or less the same, namely to appraise social reality precisely, and unambiguously. It should be sufficient, relevant, efficient and necessary. Truth and reality are absolute but knowledge and reality has an asympofatic relation. Nevertheless, researchers must strive for refined knowledge by various ways. The pilot survey serves a guideline in framing a research instrument.

SOME SUGGESTED READINGS

1. Bailey KD, *Methods of Social Research*, New York, Free Press, 1987
2. Beker, Therese L *Doing Social Research,* Singapore, Mc Graw-Hill Inc 1994.
3. Chackrabarti, P "Field-work Tradition in Indian Anthropology, Past, Present and Future ", in ; Islam, Zahidul & Hasan Shafie (eds.), *Anthropology on the Move,* Dhaka, University of Dhaka, pp.81-108, 2006.
4. Cole, Stephen *Sociological Method,* Chicago,McNally College Publishing Company.1997.
5. Dooley, David *Social Research Methods,* New Delhi, Prentice Hall of India. Pvt.Ltd. 1997.
6. Goode, William J. *Methods in Social Research,* Tokyo, McGraw-Hill Book company and Paul K Hatt Inc. 1952.
7. Moser, C. A & G. Calton *Survey Methods in Social Investigation* London, Heinemann Educational Books. Ltd.
8. Neuman, Lawrence L. *Social Research Method : Qualitative & Quantitative,* Boston, Allyan and Bacon, 1997.
9. Sanders, William B at al. *The Conduct of Social Research,* New York, Halt Rinehart and Winston.
10. Selltiz, C; Marie Jahoda *Research Methods in Social Research,* New York, Holt, Rinehart & others Winston, 1959.
11. Young, P.V. *Scientific Social Survey and Research,* New Delhi, Prentice – Hall of India, 1981